Fermat's Enigma

The quest to prove Fermat's last theorem

Everything is mathematical

Fermat's Enigma

Fermat's Enigma

The quest to prove Fermat's last theorem

Albert Violant i Holz

Everything is mathematical

One of the most fantastic pleasures of mathematics is the sharing of knowledge, whether you are a teacher, a writer or a commentator... And my family has always been a willing audience, giving me so many happy times. To my parents Eugeni and Ursula; to my brothers and sisters Deborah, Daniel, Iolanda, Alejandro, Sonia, Patricia, Verónica, Clàudia, Cristina, Silvya and Eugeni; to my wife Maria Isabel; and to my daughters Laura Elisabeth and Eulàlia Elisenda.

Published by RBA Coleccionables, S.A.
c/o Hothouse Developments Ltd
91 Brick Lane, London, E1 6QL

Localisation: Windmill Books Ltd.
Photographic credits: age fotostock

ISSN: 2050-649X

Printed in Spain

Contents

Preface

The first time a person hears about the much vaunted Fermat's last theorem, their reaction is generally: "What's all the fuss about?" The statement is so simple that one feels the temptation to grab a piece of paper and try a couple of numbers, forgetting for a moment that this is one of history's most complicated problems. One of the many who fell into this trap was Andrew Wiles, a young Brit who, when he was just ten years old, became fascinated by the theorem and the story surrounding it. After tireless attempts, the boy, armed with little more than the arithmetic he had learnt at school, had no option but to give up. However in contrast to many others, Wiles went on to become a distinguished mathematician and the problem grew to consume his working life. The epic struggle of this man, a genius who obsessively committed himself to solving a single and magnificent problem, is just one thread in the beautiful and complicated tapestry of the history of Fermat's last theorem, and which is both the start and end of this book.

The first chapter takes us back to 1993, the year in which Wiles surprised the whole world with the announcement of a proof of the last theorem. The most well-known problem in the history of mathematics had finally been solved, and journalists throughout the world reported this prodigious discovery. It is a shame that shortly after, experts were to find a number of flaws in the proof, flaws that it seemed could be solved rapidly. However, the months passed, and Wiles – under the watchful eye of the mathematical world – remained silent. Had it all been a beautiful illusion? Had the theorem thwarted yet another attempt to unravel its mysteries, as it had done for the previous three centuries?

In the second chapter, we leave Wiles behind for a moment in order to travel back more than 3,000 years and learn about the fascinating mathematics of the Sumerians and the Indians. Fermat's conjecture is closely related to an even more famous bit of maths, the Pythagoras theorem, a key result in geometry traditionally attributed to the Greek mathematician, but actually already known many centuries earlier in other mathematical traditions in the Middle East and Asia.

The fourth chapter is a sketch of the biography of our main character, Pierre de Fermat. He was a lawyer by profession and mathematician by vocation; in Fermat's time there were no scientific publications. Mathematical discoveries were made by individuals and were shared by means of a close-knit epistolary network of kindred spirits such as Fermat, Blaise Pascal, René Descartes and the Bernoulli brothers. Having placed the theorem in such an impassioned historical context, the fourth

chapter will concentrate on the genesis of the last theorem based on Fermat's readings of Diophantus' *Arithmetica*, as well as the history of the various attempts to solve the problem throughout the following three centuries until Wiles appears. It is this clotted history of distinguished names, of Gauss, the 'Prince of mathematics', Sophie Germain, the female genius who had to pass herself off as a man, from Leonard Euler to Évariste Galois, from Ernst Kummer to Yukio Taniyama and Goyo Shimura, which takes us to postwar Tokyo.

The fifth and final chapter provides an in-depth exploration of various episodes from Wiles' solitary climb to the summit of the mathematical Mount Everest that is the last theorem, a process that would bring together thousands of years of history and development in the discipline.

The pleasure of mathematics is closely related to one's knowledge of the subject. The aesthetic payoff received from understanding a mathematical marvel requires at least some effort beforehand, which makes the result all the more appreciated. Upon finally arriving at the summit, the mathematical landscape on view is comparable to the most beautiful sonata, the most exultant nature, the most ecstatic pleasure. The hope of the author is that after reading this book, the reader will discover and enjoy some immensely beautiful mathematical scenes. Some are easy to grasp, others are a little more complex but all have been written with the aim that they can be understood. More thorny subjects are left to the other titles suggested in the bibliography. Finally, this book has been written with the intention that everyone can live the adventure that the 375-odd years of effort required to reveal Fermat's enigma has meant for mathematics and humanity as a whole.

Chapter 1

Light in the Mansion of Mathematics

In a 1997 episode of the scientific program NOVA, Andrew Wiles was asked how he would define the seven years of intense, if not obsessive, research that led him to prove Fermat's last theorem, the most celebrated mystery in the history of mathematics. Wiles replied:

> "One enters the first room of a mansion and it's dark. Completely dark. One stumbles around bumping into the furniture, but gradually you learn where each piece of furniture is. Finally, after six months or so, you find the light switch, you turn it on, and suddenly it's all illuminated. You can see exactly where you are. Then you move into the next room and spend another six months in the darkness."[1]

The "darkness" that the British mathematician was talking about was the same darkness that had confused vast numbers of mathematicians and, finally, led them to give up, throughout the last three and a half centuries. The conjecture, formulated at some undetermined time in the 1630s by Frenchman Pierre de Fermat (1601-1665), stated the following:

> If n is an integer greater than 2, then there are no integers x, y and z other than 0, that fulfil the equation:

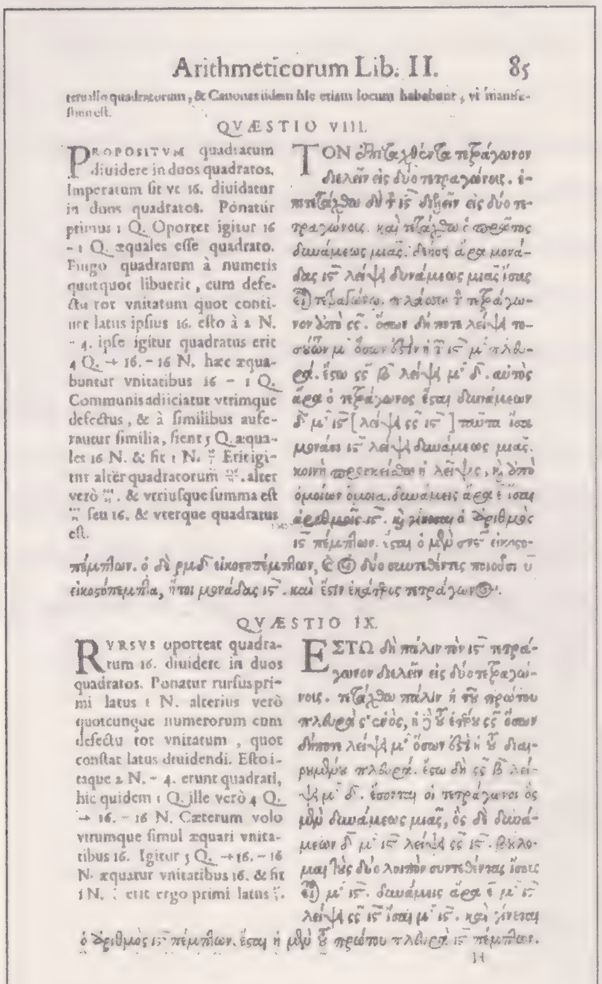
$$x^n + y^n = z^n.$$

The existence itself of the conjecture was not in the public domain until Samuel, Fermat's son, discovered it, in the form of a handwritten note in the margin of a Latin edition of Diophantus' *Aritmética*. This is not as extraordinary as it may seem,

1. The quotes of Wiles in this book, unless otherwise specified, are taken from the transcript of the aforementioned programme (*Proof*, an episode of NOVA from 28th October 1997, transmitted by the Public Broadcasting System, PBS).

given that Fermat dedicated most of his time to his profession – he was a lawyer – and only concentrated on mathematics in his spare time.

The note, as well as the presentation of the conjecture (in somewhat different terms to those quoted overleaf), contained a phrase, the promise of which would resonate like no other in the history of mathematics: "I have discovered a truly marvellous proof [of the conjecture], which this margin is too narrow to contain". So many people have wished they had been there to offer the good lawyer a blank piece of paper! Despite Samuel's best efforts, he could find no other document among his father' papers that had any reference to the promised solution except a demonstration by Fermat for the case of $n = 4$. The French genius' "truly marvellous proof" was lost forever.


Arithmeticorum Lib. II. 85

QVÆSTIO VIII.

PROPOSITVM quadratum diuidere in duos quadratos. Imperatum sit vt 16. diuidatur in duos quadratos. Ponatur primus 1 Q. Oportet igitur 16 – 1 Q. æquales esse quadrato. Fingo quadratum à numeris quotquot libuerit, cum defectu tot vnitatum quot continet latus ipsius 16. esto à 2 N. – 4. ipse igitur quadratus erit 4 Q. + 16. – 16 N. hæc æquabuntur vnitatibus 16 – 1 Q. Communis adiiciatur vtrimque defectus, & à similibus auferantur similia, fient 5 Q. æquales 16 N. & fit 1 N. 16/5. Erit igitur alter quadratorum 256/25. alter verò 144/25. & vtriusque summa est 400/25 seu 16. & vterque quadratus est.

QVÆSTIO IX.

RVRSVS oporteat quadratum 16. diuidere in duos quadratos. Ponatur rursus primi latus 1 N. alterius verò quotcunque numerorum cum defectu tot vnitatum, quot constat latus diuidendi. Esto itaque 2 N. – 4. erunt quadrati, hic quidem 1 Q. ille verò 4 Q. + 16. – 16 N. Cæterum volo vtrumque simul æquari vnitatibus 16. Igitur 5 Q. + 16. – 16 N. æquatur vnitatibus 16. & fit 1 N. 16/5 erit ergo primi latus 16/5.

Page 85 of the edition of Diophantus' Aritmética *translated by Bachet de Méziriac. The page contains problem 8 of the second book, in which the width of the margin was not sufficient for Fermat to express his "marvellous proof".*

At this point of the passionate story of the last theorem it is difficult to resist the cliché that 'the truth is often stranger than fiction'. If only Fermat had known the millions of hours of research and hundreds of thousands of pages of scientific magazines that went into seeking his promised proof, and that more than 300 years later his innocent conjecture would constitute the most well-known unsolved problem in the history of maths! 'The theory that did not fit in the margin', as it is sometimes called, has attracted a great many mathematicians without revealing its secret to any of them. Luminaries such as Carl Friedrich Gauss, Leonhard Euler, Adrien-Marie Legendre, Ernst Kummer and a long list of others tackled the problem with, we suppose, diligence and confidence in their considerable powers, but their efforts barely helped them discover partial theories for cases where n = 3, 5 and 7. Also, the further they got

into this insatiable problem, further extraordinary depths were revealed with foundations that were submerged out of sight in the algebra of number theory. The problem was of such complexity that, in the first decades of the last century, mathematicians as a whole seemed to give up on finding a solution and was at the point of relegating it to a historic curiosity. Despite the difficulties it threw up, or perhaps because of them, the legend of the last theorem breached the narrow world of mathematics and became widely known in mainstream society.

Portrait of Pierre de Fermat painted by François de Poilly as the cover illustration of the book Varia Opera Mathematica*, compiled by the mathematician's son and published in 1679.*

However, a few stubborn geniuses, who could still hear the echoes of the challenge posed by Fermat so many years ago, continued dedicating hours of obsessive work to untangling the dense network of relationships that seemed to multiply increasingly, demoting the long-awaited demonstration to a far-off and uncertain future. All of them, at some stage in their selfless task to climb this genuine Mount Everest of mathematics – perhaps after having thrown pages full of scribbles in which they had time and again found a new and fatal error into the wastepaper basket – privately asked themselves this terrible question: Could this be one of those propositions in mathematics, that "divine madness of the human spirit" as philosopher A. N. Whitehead would say, which had an answer beyond the capability of human intellectual capacity? Mathematicians such as Barry Mazur, Ken Ribet, Gerhard Frey and Gerd Faltings refused to believe it. One of the residents of the mansion haunted by the playful ghost of the French lawyer and his long-lost "marvellous proof", the rooms of which were still cloaked in darkness, was Andrew Wiles. Wiles was a mathematician barely recognised beyond a small world of specialists, who became even more obscure after withdrawing as a virtual recluse in 1986 to work on a top-secret problem.

Monday, Tuesday...

In June 1993, the Department of Pure Mathematics at the University of Cambridge, under the direction of Australian John H. Coates, called an international conference on the so-called Iwasawa theorem, an area of study within the theory of numbers regarding semi-elliptical curves, an undoubtedly intimidating term but one which is also of great importance in our story. Among the speakers was an old student of Coates, who had worked with him in proving a partial case of the famous Swinnerton-Dwyer conjecture – the Clay Institute have offered one million dollars to anyone who can provide a proof). The speaker had reclusive tendencies, but was nevertheless an excellent mathematician who had quit Cambridge before he had even graduated to move to the prestigious US University of Princeton. Despite Cambridge being an institution that celebrated being a collection of individual researchers, some with highly individual personalities, it had been a long time – perhaps seven years – since anyone had heard from Coates, to the extent that it seemed like the ground had swallowed him up. Actually, those seven years coincide, although Coates has not made this plain, with the time that had passed since American Ken Ribet had proved the so-called 'epsilon conjecture', postulated by Frenchman Jean-Pierre Serre on a masterful intuition by German Gerhard Frey. This suggested to the specialists that the legendary Fermat's last theorem was closely related to an important result from the 1950s, the Taniyama-Shimura conjecture, which, in turn, concerned certain properties of semi-elliptical curves.

After so many years out of the spotlight, it was a pleasant surprise that the ex-student in question, who was none other than Andrew Wiles, accepted the invitation. The surprise became plain curiosity when, asked by Coates how much time he would like reserved for his talk, Wiles, a painfully shy man with a notable aversion to public speaking, requested no less than three hours. Coates, who was naturally curious, asked him what important development deserved a three hour presentation, to which Wiles responded, as he had done to other associates who were attending the conference: "Come and see for yourself."

The formal name of this mysterious presentation, "Modular Forms, Elliptical Curves, and Galois Representations", listed a series of well-known mathematical terms, but Wiles kept quiet over the relationship between them, behaving like a man in possession of an unspeakable secret. This discretion was quite unusual, even in the case of professional mathematicians, a group of generally gregarious but reserved people who tend to keep their colleagues up to date on their research, even if only in search of a bit of advice.

So expectation was high when Wiles entered the conference hall of the Augusta Institute armed with notes on Monday. Under the watchful eye of the twenty-odd attendees, this willowy, tall and gaunt man of barely forty years whose slight baldness gave him that *egg-head* look that is so typical of scientists, reeled off several partial results of great scientific interest. Word spread quickly and for the next session, scheduled for the Tuesday, the classroom was packed. Once the second presentation had finished, and the mathematician had left in respectful silence, the attendees exploded into animated chatter, electrified by what they had just witnessed. By now it was clear that Wiles was not only attempting to recite a mere sequence of results, as notable as they were, but his speech as a whole was leading up to something; in fact, it constituted a carefully constructed argument of enormous complexity whose conclusion was... what? Ken Ribet, who was one of the attendess, had no doubt about it. "Wiles' presentation could only have one climax, one final objective,"[2] he said later.

...and Wednesday

On 23 June, the hallway leading to the classroom where Wiles had to give the third and last of his speeches was very crowded. Some of the attendees had taken cameras to the event, and they did not hesitate in taking numerous photos of the willowy mathematician, who was contemplating the expectation generated around him with unnatural calm. When he finally got a little bit of silence, Wiles began the third hour of his presentation, one of the most significant in the recent history of science. One notation on the blackboard followed another, and the expectation of the attendees grew and grew. Finally, the chalk in Wiles' hands drew the last lines of some symbols, which, despite being formulated in more contemporary mathematical terms, substantially expressed the same as the enigmatic note that a certain French mathematician had written in the margin of a Latin book more than three centuries before. "And this proves Fermat's last theorem," Wiles said. "I think I will stop there."

In a dark old mansion the last light came on, and the ghost that haunted it was finally exorcised.

2 Free translation of the quote taken from Amir D. Aczel's *Fermat's Last Theorem*. The content of this chapter owes much to this already classic study of Fermat's theorem and its resolution.

A mathematician in the headlines

Word of Wiles' exploits spread through the mathematical community with so much fanfare that mainstream society began to pay attention. The front page of the *New York Times* on 24 June had the following headline: "At Last, Shout of 'Eureka!' in Age-Old Math Mystery", and most of the big dailies across the world dedicated similarly privileged spaces to the news. The extraordinary story of Fermat and his last theorem was the subject of special in-depth television programmes, and

Mathematician Andrew Wiles during his presentation on 23 June, 1993, in which he demonstrated Fermat's last theorem.

practically overnight the shy, reclusive Wiles had to get used to his celebrity status. "After seven years, solving the problem gave me a marvellous feeling," the British mathematician remembered in 1997. "I had finally done it." However, the quote continues: "It was not until later that I knew there was a problem."

Effectively, what should have been a path to glory and professional recognition had become, just a few months after his dream of a presentation at Cambridge, a genuine nightmare began, the cause of which could be summarised as follows: "The old ghost would not rest."

But in order to understand exactly what happened after the ecstasy of June 1993, and to evaluate both the path taken by Wiles leading up to those magical days at Cambridge and the painful journey through the wilderness that he was compelled to take until his final and happy vindication two years later. We must go back to the root of the problem and try to understand it in its full extent and complexity. To do that we must go back thousands of years in time and embark on a fascinating individual odyssey that will take us from the dawn of mathematics, 2,000 years before Christ, to the sophisticated edifice of modern algebra and the theory of numbers. Going back to the beautiful Ithaca will ease our curiosity and satisfy our eagerness for adventure. We will thank the old ghost of Fermat and his last theorem for the journey taken and the landscapes visited, which are none other than pure human intellect in its greatest splendour.

Chapter 2

It all Began in Sumeria

Who said the history of mathematics is not important? It is a history that contains the foundations of our thoughts, the evolution of ideas and the keys to the future. It is fundamental to learning mathematics and is the best way to get a full – and dare we say pleasurable – view of the discipline. The enigmatic story of Fermat is rooted in a remote past that takes us back to Sumeria and India several thousand of years ago. The trail unerringly leads us to Pythagoras' famous equation, which states that if x and y are the sides of a right-angle triangle and z is the hypotenuse, then $x^2 + y^2 = z^2$.

Out of all the mathematicians that have ever existed, Pythagoras is doubtless the most famous, and among all the theorems that populate the field of mathematics, his theorem is the best known. So it may come as a surprise to discover that the theorem had already been discovered hundreds of years before the famous mathematician was even born. If we were to dare to change its name, what should we call the theorem instead?

Our story begins in 1800 B.C. in the area surrounding Larsa, an important Sumerian city to the south of what is now Iraq. After meticulously mixing a supply of wet clay, a scribe carefully rolls it out to create a tablet. Stylus in hand, he prepares to record a table of numbers that will survive for thousands of years.

The Plimpton 322 tablet

In 1922 George Arthur Plimpton, an editor from New York, acquired this ancient tablet from Edgar J. Banks, a dealer in archaeological objects. Although it was in relatively good condition, it had a big crack half way down the right side and the upper left part had deteriorated and was unreadable. However, perhaps most interestingly, all this suggested that the original tablet had been larger, since the left-hand side ended in a highly irregular manner, as if it had been broken off. Perhaps it was broken when it was dug up? Regardless, what has been handed down to us is a tablet approximately 13 cm wide, 9 cm high and 2 cm thick. According to Banks, the tablet comes from Senkereh, the ancient city of Larsa. Subsequent studies comparing the style of writing with other tablets from the same period have verified this origin. Similarly, researchers

The Plimpton 322 tablet.

have dated the tablet to between 1822 and 1784 B.C., four years before the capture of the city of Larsa by Hammurabi in 1762 B.C. Plimpton died in 1936, bequeathing the tablet, alongside the rest of his collection, to the University of Columbia, where it is currently held with catalogue number 322. From that point on, the tablet has been known as Plimpton 322.

The Babylonian base-60 number system

What mystery does the tablet reveal? The numbers it contains are arranged in four columns and written in base-60 using a different numbering system to our own. It is thought that this system, known as sexagesimal, has its origins in the Sumerian culture of the third millennium B.C. and was subsequently handed down to the Babylonians. It is still used today in the measurement of time, angles and geographical coordinates. In a unique example of the coexistence of decimal and sexagesimal systems, one hour is divided into 60 minutes and one minute into 60 seconds. However divisions which are smaller than seconds are now counted using the decimal system, hundredths of a second and thousandths of a second etc. Not even the mighty decimal system has been able to fully displace the sexagesimal system used by our Sumerian ancestors. A circumference continues to be divided into 360 degrees, as it has been for thousands

of years. The sundial provided the model for our analogue wrist watches for measuring time, and digital watches continue to imitate their sexagesimal design in defiance of historical events. Years and centuries are counted using the decimal system, although the days continue to have 24 hours.

Why did the Sumerians use the sexagesimal number system? It is down to some marvellous properties of the number 60. Perhaps the most interesting is the impressive list of its divisors. There are no fewer than twelve numbers that divide into 60 exactly: 1, 2, 3, 4, 5, 6, 10, 12, 15, 20, 30, 60. There is no number below 60 with so many divisors. This property of divisibility is especially attractive when working with fractions as it makes it much easier to carry out calculations. At that time, when there were still no calculators other than the human brain, such help was very useful. Many mathematicians think that the extraordinary properties of the number 60 are sufficient in their own right to justify the Sumerian's decision to use the sexagesimal number system for counting.

Other notable properties include many relationships between 60 and the prime numbers. To begin with it comes between the two twin primes (59 and 61), and is the result of adding two other twin primes (29 + 31). It can also be obtained by adding together four consecutive primes (11 + 13 + 17 + 19).

However, perhaps the most surprising property of the number 60 in this respect is that it is the smallest number that can be obtained in six different ways as the sum of two prime numbers. We can see this magnificent property in the following table.

N	The smallest number which can be expressed in *n* different ways as a sum of primes.	Ways of expressing the number as a sum of two primes.
1	4	2 + 2
2	10	3 + 7 / 5 + 5
3	22	3 + 19 / 5 + 17 / 11 + 11
4	34	3 + 31 / 5 + 29 / 11 + 23 / 17 + 17
5	48	5 + 43 / 7 + 41 / 11 + 37 / 17 + 31 / 19 + 29
6	60	7 + 53 / 13 + 47 / 17 + 43 / 19 + 41 / 23 + 37 / 29 + 31

In the 4th century, Theon of Alexandria had already suggested that the use of the sexagesimal system could be explained by the fact that 60 is the smallest number that can be divided by 1, 2, 3, 4, 5 and 6. Expanding on this idea, the mathematician J.G. van der Galiën showed that if n is a positive integer the smallest consecutive divisors of which are $\sqrt{n}$, then n can only be a prime number, double a prime number or one

of the numbers 1, 8, 12, 24 and 60. As a consequence of this result, we can conclude that 60 is the largest composite number the first divisors of which are consecutive up to $\sqrt{n}$.

HIGHLY COMPOSITE NUMBERS

Numbers with a greater number of divisors than any number smaller than it are referred to as 'highly composite numbers'. Finding a list of the first highly composite numbers is straightforward, as shown by the following table. However, there is still no formula that makes it possible to determine all these numbers.

Highly composite number	Factorisation as a product of primes	Number of divisors	List of divisors
2	2	2	1, 2
4	2^2	3	1, 2, 4
6	$2 \cdot 3$	4	1, 2, 3, 6
12	$2^2 \cdot 3$	6	1, 2, 3, 4, 6, 12
24	$2^3 \cdot 3$	8	1, 2, 3, 4, 6, 8, 12, 24
36	$2^2 \cdot 3^2$	9	1, 2, 3, 4, 6, 9, 12, 18, 36
48	$2^4 \cdot 3$	10	1, 2, 3, 4, 6, 8, 12, 16, 24, 48
60	$2^2 \cdot 3 \cdot 5$	12	1, 2, 3, 4, 5, 6, 10, 12, 15, 20, 30, 60
120	$2^3 \cdot 3 \cdot 5$	16	1, 2, 3, 4, 5, 6, 8, 10, 12, 15, 20, 24, 30, 40, 60, 120
180	$2^2 \cdot 3^2 \cdot 5$	18	1, 2, 3, 4, 5, 6, 9, 10, 12, 15, 18, 20, 30, 36, 45, 60, 90, 180
240	$2^4 \cdot 3 \cdot 5$	20	1, 2, 3, 4, 5, 6, 8, 10, 12, 15, 16, 20, 24, 30, 40, 48, 60, 80, 120, 240
360	$2^3 \cdot 3^2 \cdot 5$	24	1, 2, 3, 4, 5, 6, 8, 9, 10, 12, 15, 18, 20, 24, 30, 36, 40, 45, 60, 72, 90, 120, 180, 360
720	$2^4 \cdot 3^2 \cdot 5$	30	1, 2, 3, 4, 5, 6, 8, 9, 10, 12, 15, 16, 18, 20, 24, 30, 36, 40, 45, 48, 60, 72, 80, 90, 120, 144, 180, 240, 360, 720
840	$2^3 \cdot 3 \cdot 5 \cdot 7$	32	1, 2, 3, 4, 5, 6, 7, 8, 10, 12, 14, 15, 20, 21, 24, 28, 30, 35, 40, 42, 56, 60, 70, 84, 105, 120, 140, 168, 210, 280, 420, 840
1260	$2^2 \cdot 3^2 \cdot 5 \cdot 7$	36	1, 2, 3, 4, 5, 6, 7, 9, 10, 12, 14, 15, 18, 20, 21, 28, 30, 35, 36, 42, 45, 60, 63, 70, 84, 90, 105, 126, 140, 180, 210, 252, 315, 420, 630, 1,260
1680	$2^4 \cdot 3 \cdot 5 \cdot 7$	40	1, 2, 3, 4, 5, 6, 7, 8, 10, 12, 14, 15, 16, 20, 21, 24, 28, 30, 35, 40, 42, 48, 56, 60, 70, 80, 84, 105, 112, 120, 140, 168, 210, 240, 280, 336, 420, 560, 840, 1,680

From the decimal metric system to the sexagesimal number system

Yet these and other answers have failed to fully satisfy researchers. In fact, there is archaeological evidence that until approximately 3500 B.C., the Sumerians used the decimal metric system, Moreover, it is not known exactly what caused them to begin using the sexagesimal number system. At this point, we should note the importance of distinguishing between a number system and a metric system. The former is used for counting, adding and subtracting and, in general, for carrying out all sorts of arithmetic operations and solving numerical problems, whereas the metric system is used for measuring, lengths, surfaces, volumes, angles, weights and even time. Although it is common for both systems to be the same, this does not always have to be the case. In fact, at present we ourselves have a decimal number system that coexists alongside a sexagesimal system used for measuring time.

At the start of the last century, the German researcher Otto Neugebauer suggested the possibility that Sumerian culture first made use of a decimal metric system before switching to a sexagesimal one, of which we do not have a full record, or perhaps both were used at the same time. Neugebauer suggested that the original decimal metric system was changed to base 60 in order to make it possible to divide weights and measurements in thirds. In fact, it is known that the basic fractions of the Sumerian weights and measurement system were 1/3 and 2/3. However, it is not clear why they did not use the sexagesimal system from the outset.

Mixing populations, merging systems

Other researchers such as G. Kewitsch, have suggested that the Sumerian civilisation may have been formed as a result of two populations coming together, one of which counted in base-12 and the other in base-5. Although base-5 is not as widely used as base-10, they may share a common origin in counting using one's fingers: base-5 would make use of the fingers of one hand while base-10 would make use of the fingers of both. Continuing with this theory, as the two populations mixed and traded with each other, the base-60 system would naturally arise as the common system.

However, there are two important drawbacks to this theory. The first is that while we have evidence that base-10 was used in the area, there is no archaeological evidence to suggest that base-5 was used, although in order to overcome this inconvenience, an alternative explanation is possible. Let us suppose that shortly before 3500 B.C. there was a population that made use of a duodecimal (base-12) system which came into contact with the Sumerian people who were still using

the decimal system. The most logical explanation is that a new system progressively became standardised in order to facilitate exchanges. The ideal system would be the base that was the least common multiple between 10 and 12 to make it easy to calculate equivalences between the measurements of the two systems. That number is 60!

However, we then come up against another problem. There is no record of a population in the area that made use of a base-12 number system. By a leap of the imagination, it is possible to guess that the number 12 comes from the number of full moons in the solar year, and many units of measurement that have been handed down to us through the ages have made use of this number. For example, in the old Imperial unit system, there were 12 inches in a foot and 12 pence in a shilling. And we still count eggs in dozens! Twelve dozens make a gross, which underlines the fact that the system for counting eggs is in base-12. And 12 is also another highly composite number. However, archaeological evidence to back up theories such as this one is yet to be found.

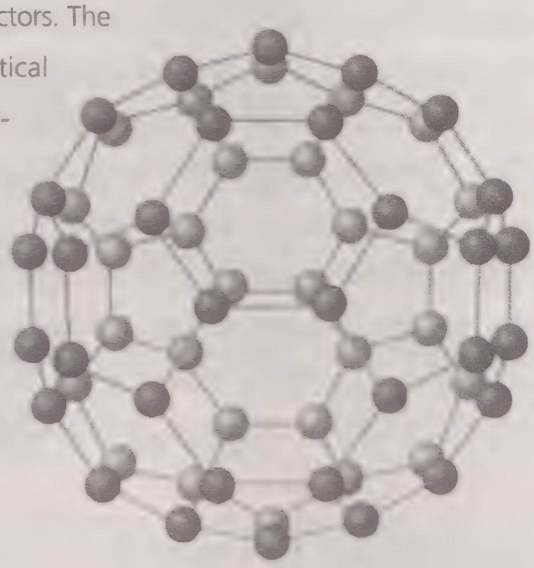
Astronomical theories and degrees

One year consists of approximately 360 days and the circumference is divided into 360 degrees. Is this a coincidence? The Sumerians were careful observers of the heavenly skies. Every night, they would see the stars of the firmament turning, and they knew that one year consisted of approximately 365 days. As such, in its journey around the Sun, each day the Earth must travel one three hundred and sixty-fifth part of the circumference. Perhaps in order to simplify their calculations, they considered that a good approximation was to divide the circumference into 360 degrees. From a mathematical perspective, the advantages were clear; however, there were also a number of geometrically interesting properties. When a hexagon is drawn inside a circle, the length of its sides is the same as the circle's radius. Thanks to this property, it is easy to draw a hexagon or divide a circumference into six equal parts using a compass. One sixth of 360 is 60. All these coincidences without doubt facilitate the task of drawing the sky and its movements and translating them into numbers. It is clear that the Sumerians were expert astronomers.

Another interesting coincidence is that the sun travels approximately 720 times its diameter in one day (the visible diameter of the Sun is 2 arc minutes), and as the Sumerian day has 12 hours, it is easy to arrive at 60. However, this would imply that the Mesopotamians had a way of measuring the visible diameter of the sun and

BUCKYBALLS

The number 60 occupies a special place in the world of geometry. In 1985 Robert F Curl, Jr., Harold W Kroto and Richard E Smalley discovered *buckminsterfullernes* (also known as *buckyballs*), measuring around one nanometre in diameter, a discovery that earned them all the Nobel Prize for Chemistry in 1996. These molecules, which are made up of 60 carbon atoms (their chemical formula is C_{60}) arranged symmetrically in hexagonal and pentagonal structures, have extraordinary properties, especially as superconductors. The superconductor *buckyballs* have the highest critical temperature that has been discovered in an organic compound and are used to make nanotubes in the field of nanotechnology. Their name is derived from the similarity of their appearance with some of US architect Richard Buckminster Fuller's lightweight domes. Their discovery rocked the scientific community since, together with graphite and diamond, they formed a new type of pure carbon.



again we lack archaeological evidence to prove this. It has even been proposed that every 60 years there is a conjunction of Jupiter and Saturn in the zodiac. There are, without doubt, many astronomical phenomena whose periodicity can be expressed using the number 60 or one of its multiples or divisors. There is good reason why it has so many!

Ways of counting

For the researchers J.J. O'Connor and E.F. Robertson the origin of the sexagesimal system lies in the way in which the Sumerian civilisation counted. According to their hypothesis, in the same way as the fingers of one's hands can be used to explain base-10 number systems, and the digits of both hands and feet can be used to explain base-20 systems, they believed that there had to be a way of counting using one's hands to arrive at the sexagesimal system. Using the thumb of the right hand and each phalanx of the remaining fingers on the same hand, starting with the pinkie, it is easy to count to 12. And to go on to count larger numbers, each time this operation has been car-

ried out, a finger of the free hand (the left) is raised until reaching 60 (12 x 5 = 60). This way of counting would also explain the common use of 12 as a reference point when it comes to counting.

Language and counting

The American Marvin A. Powell, Jr. has suggested a new theory. According to him, the sexagesimal system is rooted in an interaction between language and writing. His hypothesis is based on the appearance of base 20 in the etymology of the main Sumerian dialect, and the appearance of a base-3 system in the etymology of another Sumerian dialect. Putting both together gives the sexagesimal number system. The suggestion is that the word for sixty (*nis*) in Sumerian meant something like 'three twenties'; however, unfortunately its etymology is unknown.

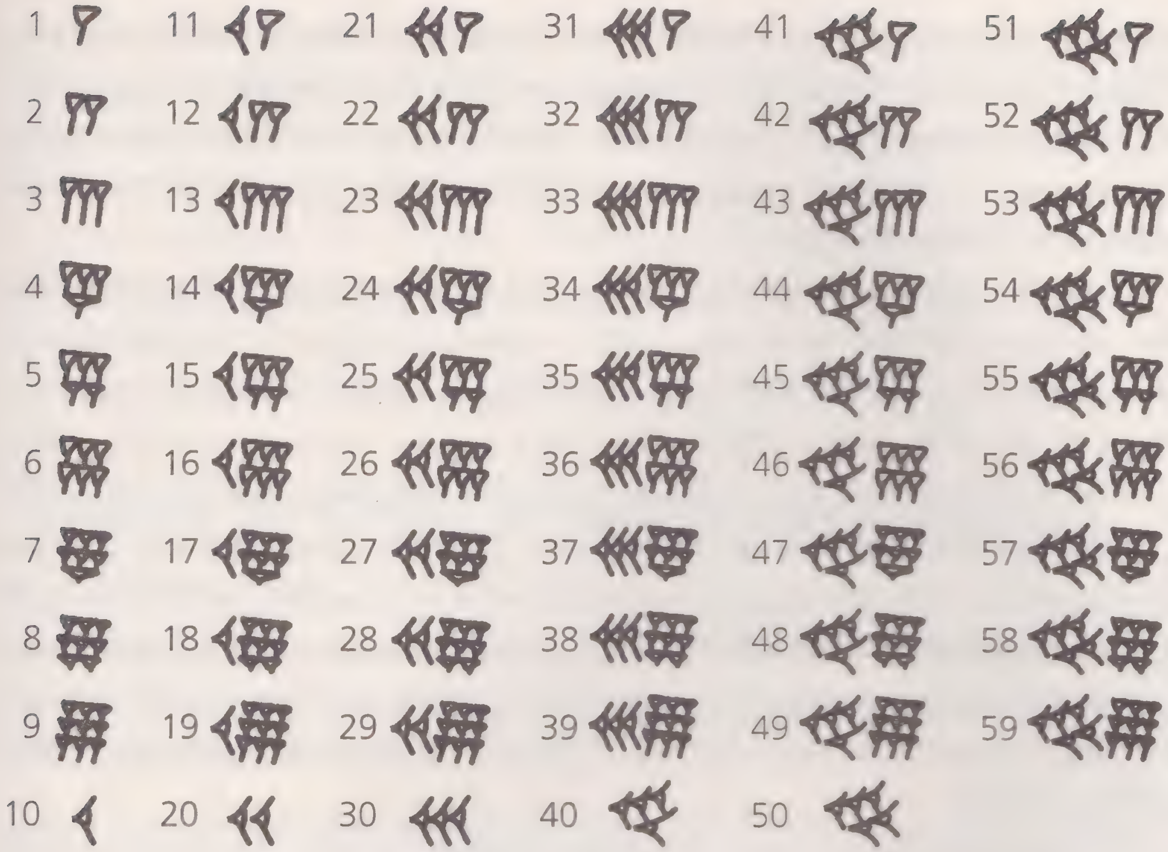
To conclude, researchers have been trying to find the origins of the sexagesimal number system for centuries, however in the end, all the theories that have been suggested are uncertain due to a lack of archaeological remains to substantiate them. We can only hope that new evidence will be found in the future.

Two symbols to count the world

One could be forgiven for thinking that the Babylonian sexagesimal number system has 60 different units, just like as there are 10 different digits in our decimal system. However, this is not the case. The Babylonians only made use of two symbols, which were used to write any number. They had an ingenious system that mixed positional notation and an additive numbering system.

Positional notation is when the numerical value of a digit differs according to the place it occupies. As an example, the decimal system, which we currently use is a positional system because a 2 placed in the first column from the right represents two units, whereas if it is placed in the second column from the right, it represents two tens.

An additive numbering system is when each symbol has a value, independent of its position in the number. Adding all the values represented by each symbol gives the value of the number.



Collection of the symbols from the Babylonian sexagesimal number system.

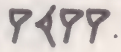
The additive system

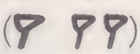
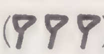
The Babylonian way of counting from 1 to 9 was extremely straightforward: Simply draw one symbol, two symbols... Thus far, their numbering system is additive. Although it is common for the symbols to be placed in specific positions, they in fact all have the same value. Each symbol has the value of one unit. As the symbols were to be carved into clay tablets using styluses, the vertical symbols were drawn in the shape of a wedge. They were grouped symmetrically, as shown in the illustration above.

A different symbol was used for the number 10, a wider wedge on its side. This made it possible to accumulate tens and units until reaching 59. As such, the system continued to be additive. Some symbols had a value of 1, whereas others had a value of 10.

The positional system

From 60 onwards, the system becomes positional. In order to represent the number 60, a symbol is written in the second position from the right. This is why the system

is referred to as sexagesimal, because a symbol in the second position has the value of 60. It is now easy to count all the way up to 3,600. As an example, the number 72 is written as follows . Starting from the right, the first two vertical symbols have a value of 1, the horizontal one a value of 10 and the other vertical one, a value of 60. In total $60 + 10 + 2 = 72$.

A problem arises when we want to write 62. Two vertical symbols placed in the first positions, followed by another vertical symbol placed in the second position (indicated by a space). The number must be written carefully in order to avoid confusing 62 () with 3 (). Moreover, if there is just one vertical symbol, how can we tell the difference between 1 and 60? Sometimes this is just impossible. The number zero had yet to be invented, and this meant that it was easy to make mistakes when it came to interpreting numbers. However, if the numbers appeared within a table instead of being isolated in a text, it was much easier to know which position was occupied by each symbol. Even so, it was necessary to take care and it must be assumed that the scribe knew what they were writing in order to avoid making mistakes.

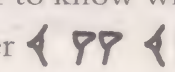
As an example, the following sexagesimal number has been translated into decimal notation:

As a first step it is necessary to read the sexagesimal number position by position and write this in decimal notation. Thus we obtain the number 20-11-1-23.

We can then calculate the value in decimal notation. On the right we have 23 units; the 1 in the second place has a value of 60; the eleven in the third position has the value of 60x60 (i.e. 60^2); and the final number in the fourth position has the value 60x60x60 (i.e. 60^3). Finally we can obtain the number in decimal notation:

$$20 \cdot 60^3 + 11 \cdot 60^2 + 1 \cdot 60 + 23 = 4,359,683.$$

Decimals

Just as there was no number for zero, sexagesimal notation did not have a decimal point. As such, it was also necessary to make use of context in order to know where the decimal point began. As an example, the sexagesimal number  is translated into decimal notation below, assuming that it is a number smaller than one unit. First of all, as above, the sexagesimal number is read position by position and

written in decimal notation. This gives the number 10-2-11 (note the space which separates the 10 from the 2 in the sexagesimal number). We can then calculate its value in decimal notation. On the left is the number 10, which represents ten sixtieths of the unit (i.e. 10/60); the 2 which follows represents sixtieths of a sixtieth of the unit (i.e. $2/60^2$), and the 11 in the third position represents sixtieths of a sixtieth of a sixtieth (i.e. $11/60^3$). Thus we can obtain the number in decimal notation:

$$10/60 + 2/60^2 + 11/60^3 = 0.167273...$$

Translating the Plimpton 322 tablet into decimal notation

Scholars followed these same steps in an attempt to discover the meaning of the numbers that appear in the Plimpton 322 tablet. The first thing they did was to number the columns and carefully transcribe their content into Arabic notation as a first step to their translation.

I	II	III	IV
59 00 15	1 59	2 49	1
56 56 58 14 50 06 15	56 07	**3 12 01**	2
55 07 41 15 33 45	1 16 41	1 50 49	3
53 10 29 32 52 16	3 31 49	5 09 01	4
48 54 01 40	1 05	1 37	5
47 06 41 40	5 19	8 01	6
43 11 56 28 26 40	38 11	59 01	7
41 33 59 03 45	13 19	20 49	8
38 33 36 36	**9 01**	12 49	9
35 10 02 28 27 24 26 40	1 22 41	2 16 01	10
33 45	45	1 15	11
29 21 54 02 15	27 59	48 49	12
27 03 45	**7 12 01**	4 49	13
25 48 51 35 06 40	29 31	53 49	14
23 13 46 40	56	**53**	15

Tablet of numbers in base-60, transcribed from the original tablet, in modern notation.

In all the tables in this chapter, the numbers in italics in the top left correspond to the illegible values on the tablet, whereas the numbers in bold in the remainder of the table correspond to supposed errors made by the scribe. These numbers are then transcribed into decimal notation, following the procedure explained overleaf in order to be able to analyse them better.

I	II	III	IV
0.983402777	119	169	1
0.949158552	3,367	**11,521**	2
0.918802126	4,601	6,649	3
0.886247906	12,709	18,541	4
0.815007716	65	97	5
0.785192901	319	481	6
0.719983676	2,291	3,541	7
0.692773437	799	1,249	8
0.642669444	541	769	9
0.586122566	4,961	8,161	10
0.5625	45	75	11
0.489416840	1,679	2,929	12
0.451041666	**25,921**	289	13
0.430238820	1,771	3,229	14
0.387160493	56	**53**	15

Translation of the table of numbers into base-10.

At first glance, it appears that this collection of numbers has no meaning. Perhaps it is simply a record of information that was lost over the course of time, like so many others. However, it is noteworthy that the fourth column, which is really the first from the right, contains the numbers 1 to 15 in a sequence, as if it were a list. On the other hand, it is also curious that the first column contains a collection of sexagesimal numbers between 0 and 1, in decreasing order. Some are complicated with a high number of sexagesimal digits, see row ten, while others are simpler, such as the eleventh row. However, at first glance, it is far from obvious that there could be any kind of connection between them.

It is at this point that the relevance of the second and third columns becomes clear, since the number that appears in the third column is always greater than that in the second, and dividing the two numbers also gives a decreasing series of numbers

between 0 and 1. This makes it possible to extend the table with an additional column, (V) based on the following formula:

$$\text{Column V} = \text{Column II} \;/\; \text{Column III}.$$

Furthermore, it is easily possible to check that squaring the second and third columns and subtracting them always gives a whole number. This in turn permits us to create another column (VI) of whole numbers derived using the formula:

$$\text{Column VI} = \text{Square root of } (\text{Column III}^2 - \text{Column II}^2).$$

Finally, bringing all this information together in a new table, it is possible to correct some of the errors in the original. For example, everything seems to indicate that there is a mistake in the second row since column V does not follow the decreasing order of the numbers and column VI does not give a whole number result. The only way to fix both errors is to enter the number 4,825 in the third column instead of 11,521. Then everything fits together once more.

I	II	III	IV	V	VI
0.983402777...	119	169	1	0.704142011...	120
0.949158552...	3,367	**4,825***	2	0.697823834...	3,456
0.918802126...	4,601	6,649	3	0.691983757...	4,800
0.886247906...	12,709	18,541	4	0.685453859...	13,500
0.815007716...	65	97	5	0.670103092...	72
0.785192901...	319	481	6	0.663201663...	360
0.719983676...	2,291	3,541	7	0.646992375...	2,700
0.692709418...*	799	1,249	8	0.639711769...	960
0.642669444...	**481***	769	9	0.625487646...	600
0.586122566...	4,961	8,161	10	0.607891189...	6,480
0.5625	45	75	11	0.6	60
0.489416840...	1,679	2,929	12	0.573233185...	2,400
0.450017361...*	**161***	289	13	0.557093425...	240
0.430238820...	1,771	3,229	14	0.548467017...	2,700
0.387160493...	56	**106***	15	0.528301886...	90

Extension of the table and the correction of errors (corrected values are indicated with an asterisk).

However a final surprise comes from the interpretation of the first column. Dividing the second column by the sixth and squaring the result gives the number in the first column up to the last decimal place. Incredible! This allows us to correct all the errors that are thought to be present in the table.

But where have all these numbers come from? By this point, it has become quite clear that it is not mere coincidence that they appear on the tablet, and explanations have been suggested by many scholars over the decades. As a first guess, perhaps the table is a roughly complete collection of Pythagorean triples (columns II, III and VI), that is to say, whole numbers that satisfy Pythagoras' theorem. Column II would correspond to the smaller leg (or cathetus), column III would appear to be the value of the hypotenuse and column VI the value of the larger leg. In fact, the Akkadian text that heads columns II and III suggests this is the case, and that column VI forms part of the 'lost' tablet.

But why select such complicated triples? There are many examples that are much simpler, such as (3, 4, 5), (6, 8, 10) and (5, 12, 13). All these correspond to whole number sides of a right-angle triangle and do not appear in the table. Although it is true that the table may have gone on to list others, it seems logical to think that among the first fifteen examples, one of the simplest would appear.

Otto Neugebauer's hypothesis

All this led the mathematician Otto Neugebauer to suggest that the numbers in columns II and III are in fact the result of calculations based on much simpler values. In 1951 Neugebauer made the claim that the person who created the table knew the formulae for generating Pythagorean triples based on these values. To do so, they first selected two relatively prime (or co-prime) natural numbers, p, q, with p greater than q. They then successively calculated the following values:

$$a = p^2 - q^2 \text{ (Column II)}$$
$$b = 2 \cdot p \cdot q \text{ (Column VI)}$$
$$c = p^2 + q^2 \text{ (Column III).}$$

It can be easily verified that in fact

$$a^2 + b^2 = (p^2 - q^2)^2 + (2 \cdot p \cdot q)^2 = p^4 - 2 \cdot p^2 \cdot q^2 + q^4 + 4 \cdot p^2 \cdot q^2 = p^4 + 2 \cdot p^2 \cdot q^2 + q^4 = (p^2 + q^2)^2 = c^2.$$

Thus these three values form a Pythagorean triple. Based on this hypothesis, Neugebauer began to extend the Sumerian table with new columns that were supposedly to be found to the left of the ones that had been recovered.

p	q	p^2	q^2	VI $2pq$	II $p^2 - q^2$	III $p^2 + q^2$	IV	p / q
12	5	144	25	120	119	169	1	2.4
64	27	4,096	729	3,456	3,367	**4,825***	2	2.370370370...
75	32	5,625	1,024	4,800	4,601	6,649	3	2.34375
125	54	15,625	2,916	13,500	12,709	18,541	4	2.3148148148...
9	4	81	16	72	65	97	5	2.25
20	9	400	81	360	319	481	6	2.222222222...
54	25	2,916	625	2,700	2,291	3,541	7	2.16
32	15	1,024	225	960	799	1,249	8	2.13333333...
25	12	625	144	600	**481***	769	9	2.083333333...
81	40	6,561	1,600	6,480	4,961	8,161	10	2.025
–	–	–	–	60	45	75	11	–
48	25	2,304	625	2,400	1,679	2,929	12	192
15	8	225	64	240	**161***	289	13	1.875
50	27	2,500	729	2,700	1,771	3,229	14	1.851851851...
9	5	81	25	90	56	**106***	15	1.8

Selection of the values of p, q *according to Otto Neugebauer (the asterisks correspond to corrected values).*

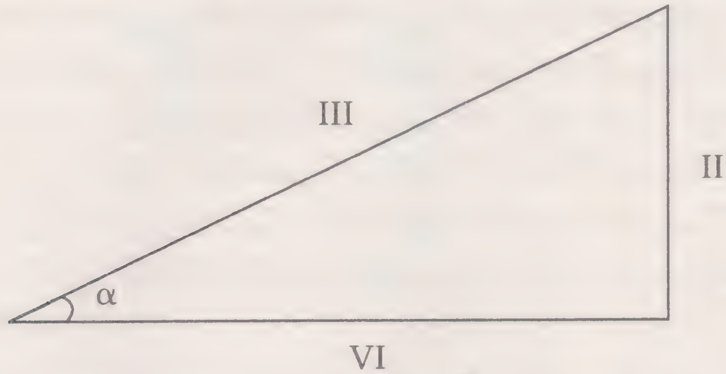
Everything appears to fit, except the eleventh row. Why is it so difficult for everything to work out properly? Why did this row fail? This row has an extraordinary peculiarity. The numbers of the Pythagorean triple (45, 60, 75) have common factors; all can be divided by 15, and upon doing so give the result (3, 4, 5), which has the associated values of $p = 2, q = 1$.

What is certain, however, is that the mystery has not yet been solved. There remain too many unanswered questions: why select these values of p, q and not others? What role did the first column play in this story?

R. Creighton Buck's explanation

In 1980, the mathematician R. Creighton Buck suggested an explanation based on trigonometry. He made scale drawings of all the right-angle triangles that appear in

the table, using the value in column II for the smaller leg and the value in VI for the larger, with the value in column III giving the hypotenuse. The angle between the largest leg and the hypotenuse is given below. It can be seen, somewhat surprisingly, that in the first triangle the legs are practically the same length and that as such, the angle formed was slightly less than 45°. From then on, the angles decrease by approximately one degree in a strictly decreasing ordering.



$$\text{Column I} = \left(\frac{\text{Column II}}{\text{Column VI}}\right)^2 = (\tan\alpha)^2.$$

The columns II, III and IV in the table make up a right-angle triangle, while column I is the result of this mathematical operation.

VII α
44.76°
44.25°
43.78°
43.27°
42.08°
41.54°
40.31°
39.77°
38.71°
37.43°
36.87°
34.97°
33.86°
33.26°
31.89°

The values of the angles of the fifteen right-angle triangles that can be formed using the numbers in columns II, III and IV of the table.

Based on all of this information, Buck ventured to suggest that column I corresponded to the square of the tangent of this angle, and that as such, Plimpton 322 showed that the trigonometric functions were already known during the period in which the tablet was made. However, this hypothesis is difficult to maintain, given that at present we have no proof from other documents from the same period that trigonometric functions were used to solve problems. All too often it is too difficult for historians to accurately define the level of knowledge of a given culture based on the documentary evidence available. There will always be a tendency to exaggerate by some and understate by others.

However the tablet remains. All the values of p, q are broken down into products of 2, 3 and 5. This means that from a mathematical perspective, the inverses of p and q always have a finite number of decimal places in the sexagesimal system. Is this the reason why these values of p, q were chosen and not others?

The interpretation of Eleanor Robson

In February 2002 the British researcher Eleanor Robson surprised the scientific community with a new interpretation of the tablet. Perhaps, after all, it is not so clear that Plimpton 322 is a list of Pythagorean triples. According to Robson, the author of the tablet could have been a mathematics teacher, who was using it as an aid in their classes in order to solve a certain type of second-degree polynomial equation. To support her hypothesis, she turned to the contents of another tablet from around the same period and place, known as YBC 6967.

This tablet provides a detailed description of a method for solving equations of the form $x-\frac{1}{x}=c$. The method works by calculating a sequence of numbers that serve as intermediate values to find the solution to the problem:

$$a_1 = c/2, a_2 = a_1^2, a_3 = 1 + a_2, a_4 = a_3^{1/2},$$

and based on these, it is easy to calculate

$$x = a_4 + a_1, 1/x = a_4 - a_1$$

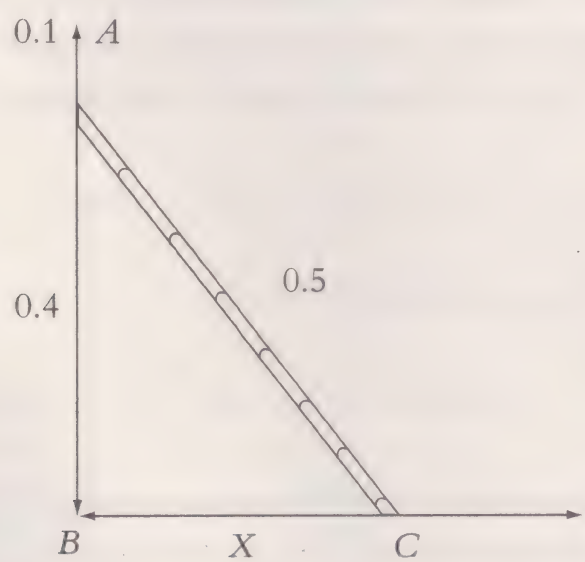
According to Robson, Plimpton 322 follows the same scheme, with a_3 in the first column, $a_1 = (x - 1/x)/2$, in the second, and $a_4 = (x + 1/x)/2$, in the third. Following this hypothesis, the values of x, $1/x$ would have been in the lost part of the table. This

would make the tablet a list of 15 exercises prepared by the teacher, with values of x, all of which were products of 2, 3 and 5, in which all the intermediate values are provided without the need to repeat the calculations each time. A genuine lesson plan, just like the ones used today by so many mathematics teachers!

Pythagoras' theorem in Sumeria

However, Robson's work brings us back to our starting point. If Plimpton 322 is not an irrefutable proof of Sumerian knowledge of Pythagoras' theorem, when is this first documented? We can look at other archaeological discoveries in the hope that new tablets provide more clues. However it is also clear, in light of our current knowledge, that Pythagoras' theorem must have appeared in one form or another throughout the long history of Mesopotamian culture, given that there is documentary evidence to show this.

There is a problem from the late Babylonian period that leaves the matter without a shadow of doubt. The text in question is as follows: "Take a cane of length 0.30, leaning against a wall. If the top slips down 0.06 from the vertical, what is the distance along the ground?" Translated into modern notation, the cane would measure 0.5 and its top would have slipped down 0.1. Drawing the problem, it becomes clear that a right-angle triangle is formed between the cane, the wall and the ground. The cane of length 0.5 would be the hypotenuse AC, with the wall AB and the ground BC forming the two legs.



The right-angle triangle formed by the Sumerian cane problem.

The same document then goes on to explain the steps to find the solution. In modern notation, the answer is given as follows:

Square 0.5 to give 0.25.
Subtract 0.1 from 0.5 to give 0.4.
Square 0.4 to give 0.16
Subtract 0.16 from 0.25 to give 0.09.
Of which number is 0.09 the square?
The answer is 0.3.
This means that the distance on the ground is 0.3.

To summarise, in order to find the value of the leg, the hypotenuse is squared, the other leg is squared and the square root of the difference is calculated. This is precisely the explanation given by Pythagoras' theorem.

There is no doubt that the author was aware that, irrespective of the length of the cane and its height on the wall, there is a general method of solving the problem. The author has also taken care in their selection of the numbers so that the problem can be easily solved in the sexagesimal number system, since all the numbers that appear in the solution can be factorised as powers of 2, 3 and 5.

And if there was still room for doubt, one of the problems that appears on numerous occasions in Babylonian mathematics is the calculation of square roots. For example, there is documentary evidence that the square root of 2 was known to an extraordinary level of precision.

All this suggests that Mesopotamian culture had an accurate knowledge of the importance of solving square roots for providing solutions to the problems of the time. It was even known that some square roots were exact, while others appeared to have an endless number of decimal places which, with patience and perseverance, mathematicians were able to approximate with increasing accuracy.

The existence of written documentation is living testament to the requirement to preserve knowledge acquired over the course of time in order to allow new generations of scholars to perfect, revise and complete the calculations. In a similar way to how astronomers left records of their observations, mathematicians also immortalised their discoveries. And this was no mean feat. It required a powerful language able to combine numbers, shapes, arguments, calculations, etc. in order to be able to transmit their discoveries to future generations.

Indian mathematics takes centre stage

In his book *The Crest of the Peacock*, the Indian mathematician George Gheverghese provides an excellent description of India's contribution to the general history of mathematics and its significant role in the specific case of the history of Pythagoras' theorem. The Indus Valley was highly significant in many respects. The fertile land made it possible for a series of dispersed communities or settlements to appear in the area in around 3000 B.C., gradually organising themselves to form cities. Mohenjo-Daro, Harappa, Taxila and Lahore are some examples of the civilisations that flourished there. The cultivation of wheat, barley, cotton and sesame, alongside the raising of livestock, provided these centres the food they required to survive. At that time, the inhabitants of those cities already felt the need to organise the land and manage the harvests, storing food in times of plenty in order to avoid going hungry in times of famine – and thus sustain a growing population. Drawing, measuring, counting, weighing… basic tasks of geometry and arithmetic were also common in the valley.

Harappa culture

Unfortunately, the texts of the Harappa culture have never been deciphered, meaning that for now, we can only base our knowledge on archaeological discoveries. These show that they had a well-established system of weights and measures. Weights have been found with uniform shapes and values, with scarcely any changes over a period of more than five hundred years. The excavations at Lothal have made it possible to classify the weights into decimal measurement standards based on a mixture of base-5 and base-10. In this manner, taking the 27.584 gram (just under an ounce) weight as the basic element, and giving this standard a value of 1, it has been determined that the scale of values was made up of weights of 0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 50, 100, 200, and 500 times that of the standard. Scales and instruments for measuring lengths have also been found. The standard measurement, referred to as the 'Indus inch' is approximately 33.5 millimetres.

The architecture of their buildings follows highly uniform guidelines, making them remarkably sound in structure. The culture also had an interest in geometry. Stone was scarce, however they discovered the useful properties of baked mud. Over the course of the years, thousands of bricks were produced by their ovens. Brick, as opposed to mud dried in the sun, withstood the rains and floods easily, an essential attribute during the river's annual inundation of arable land. Many of these bricks

have withstood the passage of time. They were of such an extraordinary quality that it is said that in the 19th century, the engineer William Brunton excavated the ruins of Harappa in order to find bricks to use as aggregate in the railway line between Multan and Lahore, some 150 kilometres long. Although fifteen different sizes of bricks have been identified, all have the proportion 4:2:1, which is still considered as optimal for construction to this very day. Arithmetic, geometry, numbers and shapes all formed part of the art, science and technology of Harappa culture.

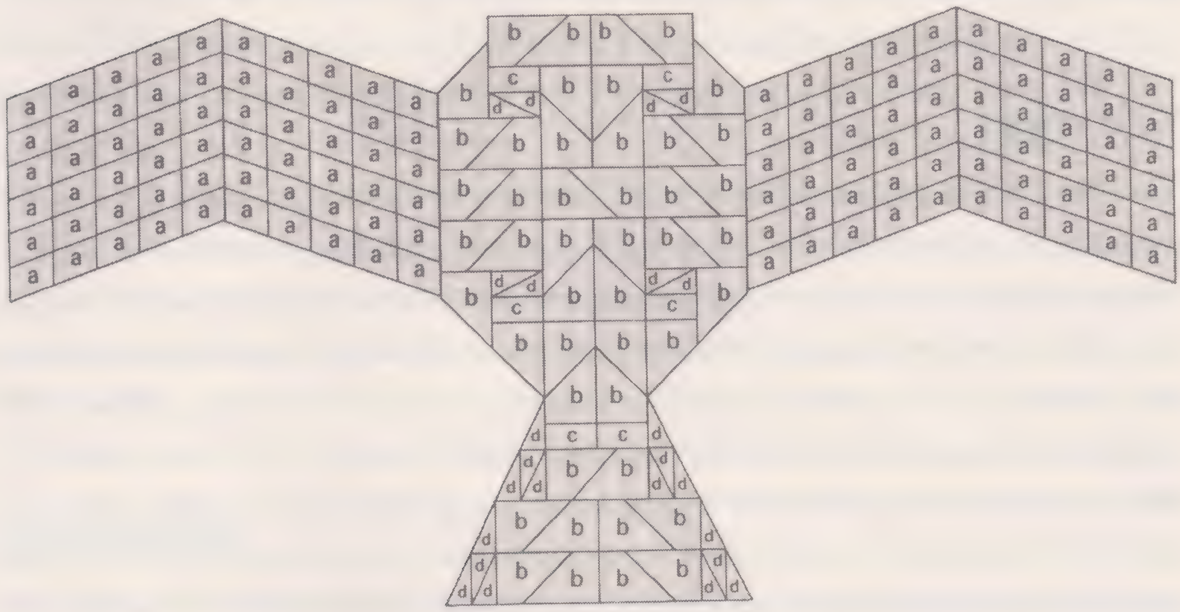
Vedic culture

Towards 1500 B.C., a pastoralist culture from the north invaded Harappa, adopting a number of its customs at the same time. The sheep-herding pastoralists, known as Aryans, spoke an Indo-European language, Sanskrit. Humankind's oldest documents are written in this language, covering philosophy, astronomy, mathematics, grammar, religion... all the things it was felt necessary to record in writing. Hymns and speeches, rites and ceremonies, formulas and magical sayings... and also extremely precise phonetic rules (for correct pronunciation), grammar (for writing correctly), metres (for mastering the rhythms of verse), astronomy (for determining the best days and times for sacrifices, and for calculating the time and positions of the Sun and the Moon on the various *naksatras* – zodiac signs) and mathematics (for determining the shapes and areas of *vedi* – altars – and the position of the *agni* – sacred fires – to ensure they were effective sacrificial instruments). Here, once again, we find documents describing Pythagoras' theorem... probably centuries before the Greek mathematician was even born.

Sulbasutras and altars

The most important mathematical sources from Vedic culture are to be found in the *Sulbasutras*, which are written in *sutras*. *Sutras* are a peculiar form of verse based on maximum brevity in order to express the essence of the idea to be transmitted. Their use led to the establishment of precise grammatical rules of the most exquisite mathematical flavour, and their poetic style leads us to define exact rules of metre and rhythm. The *Sulbasutras* discuss the shapes and measurements that must be maintained in altars in order to favour the sacrifices made there. Square and circular altars, which were easily drawn, were sufficient for domestic rites. However, more elaborate altars were required for public worship, with combinations of rectangles,

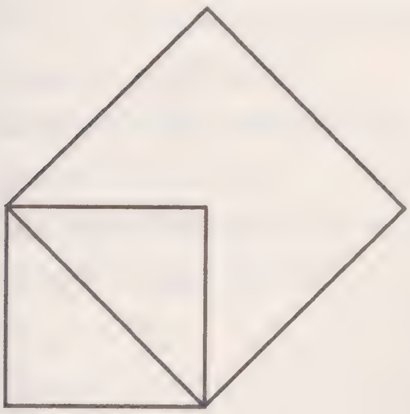
triangles, trapezoids and other geometrical forms. One of the public altars reproduced the shape of a falcon the moment before taking flight and it is thought that making a sacrifice on this alter allowed the soul of the person who made the prayer to be carried straight to heaven.



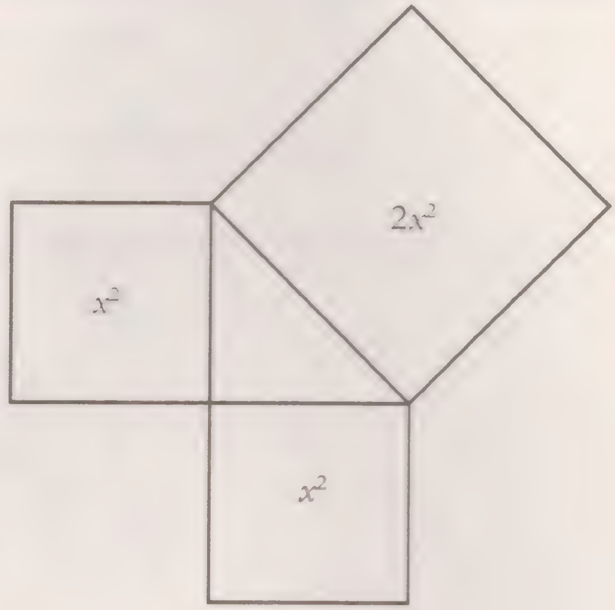
The Vedic altar shaped like a falcon; the letters refer to the different types of bricks used in its construction (source: George Gheverghese, The Crest of the Peacock*).*

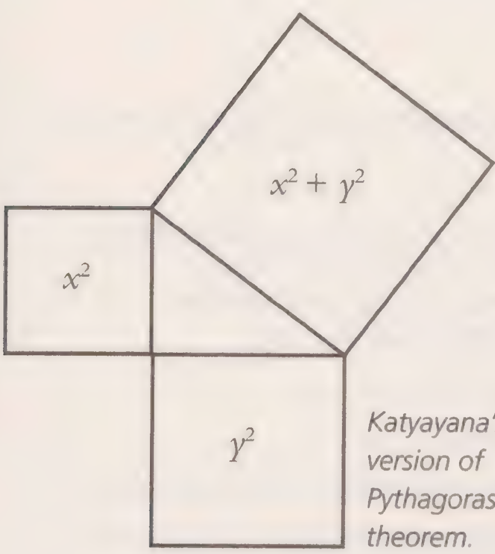
The area of an altar was one of its essential properties. Hence the importance of having formulae that made it possible to transform one geometric shape into another with the same area. The *Sulbasutras* provided solutions to all these requirements. The Baudhayana *Sulbasutra*, dating from between 800 B.C. and 600 B.C., contains a formulation of Pythagoras' theorem, a procedure for calculating the square root of 2 (correct to five decimal places), and a full series of geometric constructions. These included various ways of squaring the circle (approximately) and drawing polygons with areas that were equal to the sum or the difference of two other polygons. The meticulous precision of the shape and measurements of the altars was just as important to the rite as to the scrupulous pronunciation of the prayers (*mantras*). Later on, Apastamba wrote *Sulbasutras* explaining the topics covered by Baudhayana in more detail, and Katyayana also wrote new *Sulbasutras,* although these contained much less new material than the previous ones. Both produced their work in the century prior to Panini, the Sanskrit grammarian from the 4th century B.C.

Baudhayana explicitly established a statement of Pythagoras' theorem: "A rope (*sulba*) that is stretched in the diagonal of a square produces an area double the size of the original square." Katyayana gives a more general version: "A rope [stretched along the length] of the diagonal of a rectangle gives an [area] produced by joining the horizontal and vertical sides."



Baudhayana's version of Pythagoras' theorem. The squared area above the diagonal is double that of the original square.



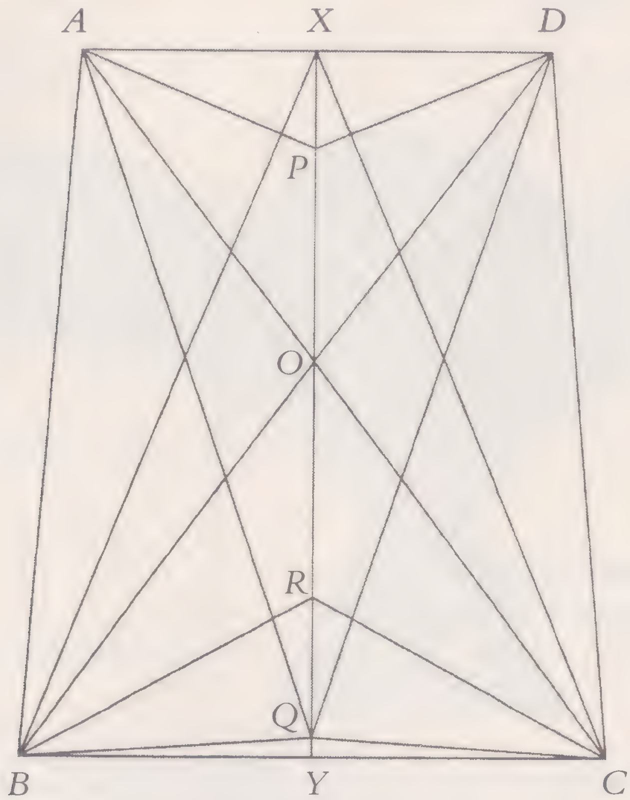


Katyayana's version of Pythagoras' theorem.

This knowledge made it possible to draw the Vedic altars with extraordinary precision. As an example, let us consider the altar referred to as *smasana*, which was used to offer *soma*, an intoxicating spirit, to the gods. The base was required to have accurate dimensions so as to ensure the sacrifice took effect, and the Apastamba *Sulbasutra* gives precise instructions as to how it should be created. These instructions, in modern notation and in the words of George Gheverghese, are as follows:

> Using a rope, mark *XY*, which is exactly 36 *padas*.
> Along this line, add points *P*, *Q* and *R*, such that *XP*, *XR* and *XQ* are equal to 5, 28 and 35 *padas*, respectively.
> Draw perpendiculars to *X* and *Y*.

Use the fact that the triangles *APX*, *DPX*, *BRY* and *CRY* are right-angle triangles with whole number sides in order to find the points *A*, *B*, *C* and *D*. In other words, make *AXD* 24 *padas* and *BYC* 30 *padas*. If the construction is correct, they should intersect at point *O* on *XY*.



AX = *XD* = 12 *padas*
BY = *YC* = 15 *padas*
XP = 5 *padas*
PR = 23 *padas*
RQ = 7 *padas*
QY = 1 *pada*
XY = 36 *padas*.

Measurements for the smasana altar. *(Source: George Gheverghese,* The Crest of the Peacock.*)*

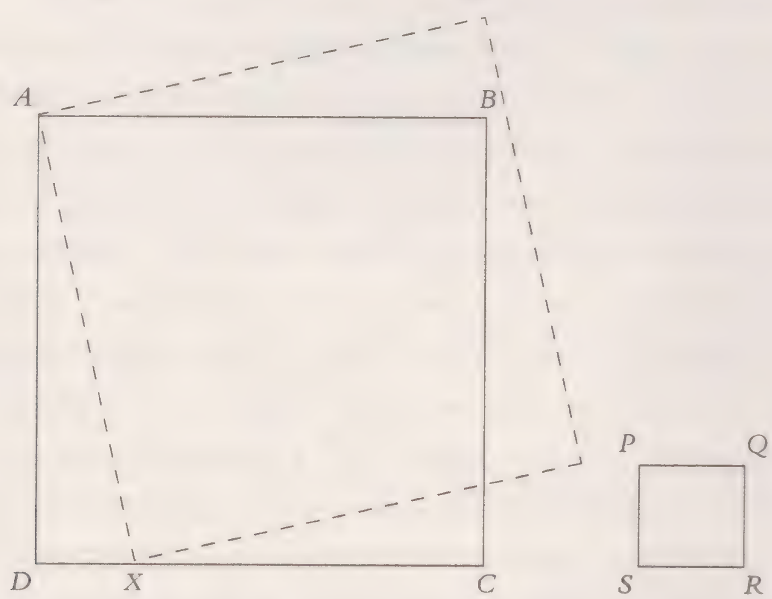
Following these directions, we obtain a collection of Pythagorean triples:

ΔAPX and ΔDPX with sides 5, 12, 13.
ΔAOX and ΔDOX with sides 12, 16, 20.
ΔAQX and ΔDQX with sides 12, 35, 37.
ΔBRY and ΔCRY with sides 8, 15, 17.
ΔBOY and ΔCOY with sides 15, 20, 25.
ΔBXY and ΔCXY with sides 15, 36, 39.

Given that the sides of the triangles are whole numbers, it is possible to draw them with extraordinary precision. However, as if this were not enough, the construction itself abounds in additional Pythagorean triples, which made it possible to

work with even greater precision. Without a doubt, both extraordinary and beautiful. Indeed, there were many other triples used in the design of a countless number of altars making it clear that the Vedic people had perfectly mastered and applied Pythagoras' theorem. For them, it was normal to solve problems that required the "merger of two equal or unequal squares in order to obtain a third". In this way, they were able to obtain an altar that was the sum of the other two. The *Sulbasutras* provide the solution. In modern notation, we have:

> Let $ABCD$ and $PQRS$ be the squares we would like to merge.
> Let $DX = SR$.
> Then the area of the square on AX is the sum of the areas of the squares $ABCD$ and $PQRS$.



This illustration clearly shows the process described in the text and is evidently along the same lines as Pythagoras' theorem, $AD^2 + SR^2 = AX^2$. (Source: George Gheverghese, The Crest of the Peacock.*)*

It is indisputable that the human mind has been aware of the beauty of geometry and arithmetic from time immemorial. From the outset, it must have been captivated by straight-line and curved shapes. Within this world, it is clear that right-angle triangles soon came to play a privileged role. A rectangle divided along the diagonal forms two right-angle triangles. Within arithmetic, the natural numbers, which are used to count, must have also occupied a privileged space. One day, it was discovered that it was possible to draw right-angle triangles with three whole number sides.

The discovery that the sum of the areas on the sides was equal to the area on the hypotenuse must have been something very special indeed.

A precious and remarkable property of a precious and remarkable shape had been discovered; the beauty inherent to right-angle triangles was worth handing down to future generations, until Pythagoras, on one of his journeys to Egypt or Mesopotamia, discovered and was captivated by it, just as we continue to be captivated to this day. And he went on to provide a proof; perhaps the first, or perhaps not. Yet either way, he fell in love with the beauty of numbers and shapes and came to declare that the world is mathematical. The exact date and the discovery of this theorem are, for the time being, a mystery, and it is highly likely that there will be no exact date nor a single discoverer. Perhaps it has been discovered a number of times by various people from various cultures. However there it is, the theorem of the right-angle triangle. Perhaps this would be the best name for such a beautiful theorem.

Chapter 3

Fermat, a Lawyer to be Reckoned With

I, who do not profess to be a mathematician, but who, whenever there is leisure, delight in mathematical studies...
François Viète

The theory of the right-angle triangle affirms that the sum of the squares of the shorter sides equals the square of the hypotenuse. In modern notation we would say $x^2 + y^2 = z^2$. As we saw in the previous chapter, integer solutions to this equation have been known since ancient times, solutions that would later become known as 'Pythagorean triples'. With these solutions, right-angle triangles with integer sides could be drawn easily. With time, formulae were even developed for finding all the Pythagorean triples, and it turned out there was an infinite number.

It occurred to Fermat to change the power of 2 for another power: 3, 4, 5... It really does not seem like such a big change, one that perhaps anyone could have thought of. However, thousands of years had passed since the theorem of the right-angle triangle was discovered when Fermat came up with this modification. The surprising thing is, not for want of trying, he could not even find one integer solution for any of the powers. Incredible in itself. He must have spent hours until he finally realised, effectively, that there was none, and that he could prove it. So he declared his theory and rested, waiting for another day until writing it down. Who was Fermat? How did this idea occur to him? What was the demonstration he had thought of? Why did he not write it down?

Fermat's life is veiled in mystery. He was an isolated mathematician whose only means of communication with his friends and passing on his findings was to write letters, a lover of mathematics who discovered surprising and marvellous things, a genius who enjoyed setting challenges for the other mathematicians of the era, to which he said he knew the solutions, but rarely published them. Much is said about him, but in reality we know a lot less about him than we would like to.

This sculpture of Pierre de Fermat, made in 1898 by Charle Barrau, is on exhibition at the Salle des Illustres del Capitolio de Toulouse. (Photograph: GFDL Nicolas Guerin.)

Place of birth, family and education

The first enigma surrounding Fermat is that it is not known for certain when he was born. His date of birth is often stated as being 20 August 1601, in Beaumont de Lomagne, in the department of Tarn et Garonne in the south of France. This date, proposed in 1844 by Louis Taupiac, supposes that Fermat was the son of Dominique Fermat, bourgeois second consul of Beaumont, and Claire de Long, daughter of Clément de Long, Lord of Barés. However, according to the death certificate that is held in the archives of the Fermat villa, he died in Castres on 12 January 1665 and the epitaph on show in the Musée des Augustins in Toulouse states that he died at the age of 57. Evidently something does not add up, because this would make Fermat's year of birth 1608.

In 1980 Abbot Dugros proposed another theory. According to him, Fermat was the son of Dominique Fermat and Françoise Cazeneuve, the daughter of an affluent art dealer from the outskirts of Beaumont. His hypothesis is based on the fact

that in 1603 Dominique married Françoise, but in 1607 he was married to Claire de Long.

In 2002 Pierre Gairin returned to the problem and dedicated his time to a thorough study of Fermat's family tree in search of a convincing answer. New data was then found that once again left the question open. Still awaiting new evidence, Fermat could have been the son of Claire de Long, the daughter of a noble French family, and could have been born around 1605 or, more probably, 1606. Even so, his epitaph still lacks a clear explanation.

It seems that Fermat had one brother and two sisters, and that he received an education from the Franciscan Friars of Grandselve from an early age. But little more is known of his childhood. However, during his whole life he felt a great attachment to the place where he was born, refusing to be impressed by the romance of Paris.

The second enigma that surrounds Fermat is that of his education. His thirst for knowledge is beyond all doubt. He translated the classics, he had an excellent command of Latin and Greek, and his mathematical knowledge was unarguable. But where did he learn mathematics? Did he learn it at university? Did a friend teach him? There are doubts about Fermat's education and professional career, although some things can be deduced from a few comments in his letters. In his youth he attended the University of Toulouse to complete his studies. His intention, and surely that of his family, was to study law, a career of great prestige that opened many doors after solid training, and offered the chance to become a court official.

Mathematical circles

Subsequently, Fermat transferred to Bordeaux, where, according to some biographies, he worked as a lawyer and where, probably in August 1626, he met Jean Beaugrand, who belonged to a circle of prestigious mathematicians who regularly met there. At the beginning of the 17th century, the scientific community realised the benefits of sharing its ideas, and begun to form mathematical networks that maintained regular communication. Questions were put forward, solutions were proposed, methods were explained, new thoughts were presented, new challenges were set... and Bordeaux was one of the centres in which those ideas bubbled over, talent was stimulated and large doses of creativity were on show. This environment undoubtedly attracted the attention of young Fermat.

Beaugrand always showed his pride among his friends at having discovered Fermat, and he was one of the people who contributed most to making known

his works. Beaugrand ended up forming part of the political class in Paris, and was widely recognised for his mathematical abilities. It is even said that he was a student of Viète. In Paris he frequented a group of mathematicians who met with Marin Mersenne, to whom he always explained that he mediated his own findings and those of Fermat. And to make oneself known to Mersenne was to make oneself known to the whole mathematical world at that time. Beaugrand also travelled to Italy, invited by Bellièvre, the French ambassador. Once there, like any mathematical tourist, he went to visit Cavalieri in Bologna, Castelli in Rome and Galileo in Arcetri, near Florence. All of them maintained a fruitful correspondence after returning to Paris. Inevitably, he told them all about his friend Fermat.

Portrait of Galileo Galilei, painted by Justus Sustermans in 1636. The Italian scientist knew of Fermat's work through a mutual friend, Jean Beaugrand.

FRANÇOIS VIÈTE

Born in 1540 in Fontenay-le-Comte, François Viète was a lawyer by trade, as was Fermat, but his passion was mathematics. When he became the tutor of a young girl named Catherine of Parthenay, he begun to study astronomy, a subject which she was interested in. Upon entering into the service of king Henry IV of France he deciphered the code used by Felipe II's troops in their communications, breaking a key of more than 500 symbols which mathematicians of the Spanish court considered impregnable. When the Spanish found out, Felipe II sent a complaint to Pope Pius V accusing Henry IV of using black magic to combat his armies.

It is said that Viète's power of concentration was such that he would often study for three days on end without even eating or sleeping. One of the most famous anecdotes about him was told by Tallemant des Réaux: "In the times of Henry the fourth, a Dutchman called Adrianus Romanus, a learned mathematician, but not so good as he believed, published a treatise in which he proposed a question to all the mathematicians of Europe, in one part it even listed all the mathematicians in Europe, but the list did not include any Frenchmen. Shortly after, a state ambassador came to the King at Fontainebleau. The King took pleasure in showing him all the sights, and he said people there were excellent in every profession in his kingdom. 'But, Sire,' said the ambassador, 'you have no mathematician, according to Adrianus Romanus, who didn't mention any in his catalogue.' 'Yes, we have,' said the King. 'I have an excellent man. Go and seek Monsieur Viète,' he ordered. Viète, who was at Fontainebleau, came at once. The ambassador sent for the book from Adrianus Romanus and showed the problem to Viète, who had arrived in the gallery, and before the King came out, he had already written two solutions with a pencil. By the evening he had sent many other solutions to the ambassador, adding that he would provide as many as he wanted, as it was one of those problems with infinite solutions." Adrianus Romanus was asking to resolve an equation of 45 degrees in which Viète immediately recognised an underlying trigonometric ratio. From there he went on to determine the other 22 positive solutions, the only ones admissible at that time. In 1595 Viète published his response to the problem set by Adrianus Romanus and to prove his mathematical expertise he concluded by proposing finding the solution to a problem of Apollonius, namely to find a circle tangent to three given circles, to which he had an answer.

In Bordeaux, Fermat also mixed with D'Espagnet, Philon and Prades. Speaking with them broadened his mathematical horizons, and he met numerous figures from the world of science. With help from Étienne d'Espagnet, advocate at the Bordeaux parliament, he familiarised himself with the works and notations of Viète. In fact, D'Espagnet's father was the first president of the Bordeaux parliament and a friend of Viète's. It should be taken into account that in those years the transfer of knowledge was heavily dependant on one's friendships and contacts.

Bordeaux was also the first city that saw Fermat's first mathematical publications appear and where he produced his first works, such as the restoration of Apollonius' Plane Loci (*Apollonii Pergaei Libri Duo de Locis Planis Restituti)*, the method of maximums and minimums (*Methodus ad Disquirendam Maximam et Minimam et de Tangentibus Linearum Vurvarum*) and some research on the magic squares.

Political and administrative career

After his stay in Bordeaux, Fermat moved to Orléans, where he finished his law studies, and finally to Toulouse, where he began a brilliant professional career. On 14 May 1631, he was named advocate at the Parliament of Toulouse and the Higher Chamber, in which public matters were debated. On 30 December, 1634, he was named advocate of the Summary Court, and in 1638 he attained the position of advocate of the Court. In August 1648 he was named advocate of the Chamber of Edit and in 1654 read his first report in the Great Chamber. In summary, Fermat had a fully fledged career in law and held positions of significant responsibility. As a result of those positions, he was granted the right to change his name to the more distinguished Pierre de Fermat.

THE FERMAT FAMILY

A month and a half after he was first named as a government comissioner, Fermat married a cousin of his mother, Louise de Long, on 1 June 1631. The couple had five children: Clément-Samuel, who was a magistrate like Fermat himself and who published much of his father's work; Jean, who was archdeacon of Fimarens; Claire, who married and whose grandson, Jean Gaillard, succeeded Jean Françoise, Samuel's son, as a counsellor; and Catherine and Louise, both dedicated to religion.

It is possible that his career was partly favoured by the plague epidemic that struck the region around 1650, claiming a great many victims. So many deaths in such a short time must have left a lot of vacancies in the magistracy, where experience was of great value and, therefore, positions were normally filled by older people. In fact, the disarray was such that in 1653 even Fermat's death was erroneously reported. He was said to be another victim of the epidemic; this information was naturally corrected later.

There are conflicting opinions regarding the diligence with which Fermat carried out his work. On the one hand, in 1664, the lawyer Pierre Saporta stated: "I would say of your judgement in issues of the Palace, where you have spent most of your life, and where you seem to have performed with great integrity, and such sufficiency in the administration of justice, that it is amazing, that having acquired the qualities of a great judge, you have been able to acquire a complete comprehension of so many other things, which are so different to this field." It should be noted that Saporta was a friend of Fermat's and that he dedicated to him his French translation of Torricelli's paper on the movement of water, published in the same year. On the other hand, in 1663, Jean Baptiste Colbert said: "Parliament

Jean Baptiste Colbert shown in a portrait by Claude Lefebvre in 1666. The minister of finance to King Louis XIV, he did not give a very favourable report of Fermat's work as a magistrate.

of Toulouse: Fermat, a man of great erudition, has contact with men of learning everywhere. But he is rather preoccupied, he does not report cases well." This document formed part of a secret report on the judiciary drawn up by Colbert, at that time Minister of Finance to King Louis XIV, who was generally seen to be quite hard on all judges. Taking into account that Fermat did not form part of the circle of friends of Prime Minister Gaspard de Fiubet, it is not surprising that this report was not all that complimentary.

In any case, it seems to be true that Fermat was reliable in his work and, as far as possible, tried to avoid getting into trouble. It was not easy work, given that at that time France was going through a period of political upheaval. On the one hand, there were tensions between the Catholics and the Huguenots; and on the other, constant battles for power. It is the period of Cardinal Richelieu and Cardinal Mazarino, and the period of the powerful musketeers. It was easy to take one side one day and fall at the hands of the other the next. The most difficult was to maintain a balance between the two. And that is surely what guided Fermat's activities. He had no ambitions for power and perhaps that is why he avoided going to Paris and focused his activities on smaller cities. He was seldom far from his childhood home

Cardinal Richelieu at the site of La Rochelle in Henri Motte's oil painting from 1881. The siege of the Huguenot fortress took place in 1628 and is one of the many examples that illustrate the turbulent times Fermat lived in.

(in 1638, 1644 and 1645, 1648 and 1649, 1655 and 1656, 1663 and 1664 he was a judge in Castres, a small city a few kilometres from Beaumont de Lomagne).

On the other hand, Fermat did not tend to mix with the middle class of the time. His position as judge meant that he had to maintain prudent distance from those who some day may have to attend court to answer his questions. In a certain sense, Fermat must have been lonely. While other people in his profession dedicated their free time to writing about the court and the judiciary, he dedicated all his time to mathematics.

The 'prince of amateurs' and Pierre de Carcavy

Fermat constantly complained about the lack of time he had. In June 1640 he wrote to his friend Mersenne "I am pressured by many worries that do not leave me much time for other things..." and later, in March 1641, "the worries about lawsuits that fill our heads prevent me from being able to read the documents, which you kindly sent me, in peace." Even so, whenever he could, he found time for mathematics and shared this passion with other colleagues in the profession and other mathematicians around the world. Fermat is a clear example of the fact that it is possible to combine profession and hobbies at the highest level. In this sense he was not unique in his time. Many of his professional colleagues were also good mathematicians and made some notable discoveries. What is remarkable in the case of Fermat is the quantity and quality of his discoveries, which gave him the title 'prince of amateurs', although American author Julian Coolidge excluded him from his celebrated dictionary of mathematics amateurs, considering him to be 'professional'.

In Toulouse, Fermat met Pierre de Carcavy, advocate of the Parliament of Toulouse between 1632 and 1636 and a great lover of mathematics. This friendship played a fundamental role in Fermat's mathematical life. Although Carcavy did not received any specific university training, and his mathematical results were not particularly relevant, the correspondence he maintained with many of the period's scientists and its role in the spread of knowledge is worth mentioning. Carcavy maintained diligent correspondence with Fermat, Huygens, Pascal, Descartes, Mersenne, Galileo and Torricelli – among many others – throughout his life

It should be taken into account that many discoveries never reached publication and that the only way of discovering them was by means of letters written by their finders. People who created new theories and methods were very important for the advance of mathematics, but those who dedicated much of their time to spreading

Portrait of Descartes painted by Frans Hals in 1649. One of the most important philosophers of the 17th century, he was always very interested in mathematics and maintained correspondence with numerous figures from his time.

word of them were as well. These letters do not just pass on results, but contain comments and suggestions of new ideas, which were later taken into account and included in other work. It must have been so exciting living at a time when each letter received could have contained a barrage of new ideas and an incentive for getting back to work. The letters however, could take weeks to get to their destination, that is if they were not lost on the way. There was a different rhythm to life, the speed of the postal service was nothing like an email.

In 1636 Carcavy bought a advocate's office at the Great Council of Paris and moved to the capital. Fermat continued his correspondence with him and sent him a lot of his work. Once in Paris, Carcavy also met Mersenne, a key figure in the history of mathematics and science, face to face.

Marin Mersenne

Born in 1588 to a poor family, at the age of sixteen Marin Mersenne signed up to La Flèche, a Jesuit school that sought to be a model in the education of children, regardless of their families' resources. It was also the school of Descartes, with whom Mersenne later had a close relationship. He later went to Paris and joined the Collège Royal du France to study philosophy and the Sorbone to study theology, finishing his studies in 1611. He decided to follow a monastic life of study, joined the Order of the Minims in 1611, and a year later was ordained as a priest. Members of this order were characterised by their austerity and also their exceptional academic education, they were known as *les bonnes hommes* due to their kindness. Mersenne's initial work focused on theology, but as time passed he began to develop a growing interest in science. He firmly believed that mathematics was the basis for all sciences and

that communication was essential for the advance of science and mathematics. Scientists were very isolated from one another, and many of their discoveries became known years after their deaths, or were even lost forever. Mersenne realised that far more progress could be made if people worked together. In around 1623 he began to create a group of scientists who met regularly in his Paris friary. At these meetings, science was discussed and results, calculation methods and new findings reported. This was the origin of the Académie Parisiensis (also referred to as Mersenne's circle).

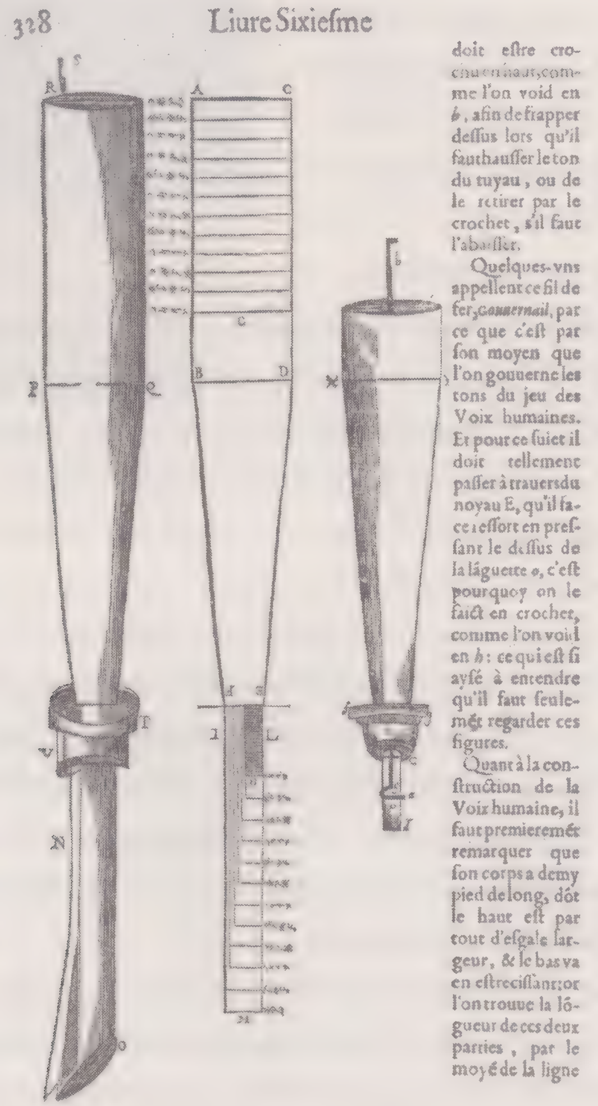
Monk and mathematician Marin Mersenne.

Not satisfied with the face-to-face meetings, Mersenne maintained a prolific correspondence with many scientists around the world, thus creating a genuine international community. The priest was up-to-date with all the discoveries of the time, and it was said that once he was informed of a finding it was as if the whole world had found out about it, as he would pass it on to all the people who were working on that subject or who might be interested in the new discovery. The list of scientists which he kept in contact with is enormous, among them Isaac Beeckman, Bernard Frénicle de Bessy, Charles Cavendish, Florianus Crusius, Girard Desargues, Descartes, Fermat, Galileo, Pierre Gassendi, Jean Baptiste van Helmont, Hobbes, Christiaan and Constantin Huygens, Claude Mydorge, Étienne and Blaise Pascal, Nicolas Claude Peiresc, John Pell, Jean Rey, André Rivet, Gilles Personne de Roberval, Martin Ruarus, Samuel Sachieri and Evangelista Torricelli. When he passed away in 1648, letters from 78 different correspondents were found in his cell.

Moreover, Mersenne made the most of the travelling he had to do by establishing contacts with emerging scientific communities and the most prominent scientists. Between 1629 and 1630 he travelled to the Netherlands, in 1644 to Provence and Italy, where he met Torricelli, and in 1646 to Bordeaux, where the Académie Royale des Sciences was later formed.

Mersenne's scientific interests encompassed the most varied fields of physics and mathematics, and his patronage served as a *leitmotif* to a large part of the period's scientific activity. The problems that Mersenne worked on covered music, optics, heat, mechanics, hydrostatics, analysis, algebra, number theory, etc. His opinions were quickly shared with other researchers and those who were keen to continue moving forward and expanding on his results. Thus, Mersenne never worked alone; reciprocally, those who maintained correspondence with Mersenne found in him an exceptional interlocutor with whom to share their ideas and complete their research.

In 1627 Mersenne published *L'Harmonie Universelle*, in the preface of which he recognised Fermat's mathematical genius. Among other things, this text establishes



328 Liure Sixieſme

doit eſtre crochué en haut, comme l'on void en *b*, afin de frapper deſſus lors qu'il faut hauſſer le ton du tuyau, ou de le retirer par le crochet, s'il faut l'abaiſſer.

Quelques-vns appellent ce fil de fer, *Gouuernail*, par ce que c'eſt par ſon moyen que l'on gouuerne les tons du jeu des Voix humaines. Et pour ce ſuiet il doit tellement paſſer à trauers du noyau E, qu'il face effort en preſſant le deſſus de la láguette *o*, c'eſt pourquoy on le faict en crochet, comme l'on void en *b*: ce qui eſt ſi ayſé à entendre qu'il faut ſeulemẽt regarder ces figures.

Quant à la conſtruction de la Voix humaine, il faut premieremẽt remarquer que ſon corps a demy pied de long, dõt le haut eſt par tout d'eſgale largeur, & le bas va en eſtreciſſant: or l'on trouue la lõgueur de ces deux parties, par le moyẽ de la ligne

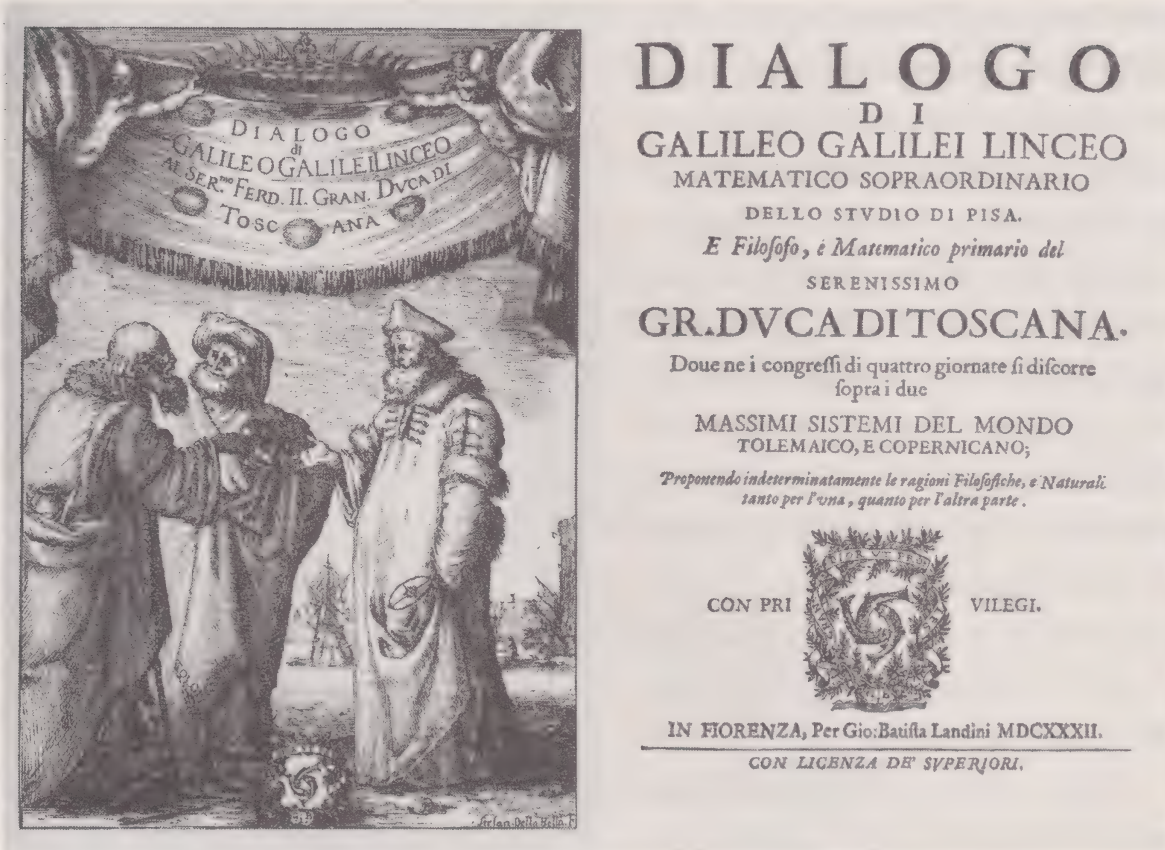
A page from Mersenne's work L'Harmonie Universelle, *which mentions the mathematical genius of Fermat in the preface.*

that the frequency with which a string vibrates is proportional to the square root of the tension, and inversely proportional to its length, its diameter and the square root of its specific weight, assuming that all other variables remain constant when one of these amounts varies. When Mersenne met Christiaan Huygens he shared his knowledge with him. As a result of this exchange Huygens later published his *Theory of Music.* In 1646 Huygens attempted to move to Paris to be nearer to his mentor, but he did not manage it until years after Mersenne's death.

Portrait of Dutch scientist Christiaan Huygens.

Galileo Galilei's ideas also interested Mersenne, who contributed enormously to spreading them across Europe. On 21 February 1632, *Dialogo Sopra i Due Massimi Sistemi del Mondo*, one of Galileo's fundamental works, in which Copernicus' heliocentric theory was defended, was printed in Florence.

This event undoubtedly upset the doctrines established by the Catholic Church of the time, and therefore, in 1633 Galileo was judged and sentenced by the Santo Oficio in Rome. As a result of the proceedings, the book was banned, but numerous copies had already been distributed around the whole of Europe and many scientists had already read it. Among them was Mersenne, who, intrigued by the free fall of bodies demonstrated in Galileo's work, decided to personally carry out a series of experiments. In 1634 Mersenne published his results, in which he confirmed the relationship between the acceleration during a fall and the square of the time taken. He also tried to resolve the question of whether the variation of the velocity of falling was continuous, as Galileo believed, or non-continuous, as Descartes believed.



Frontispiece of Galileo's work Dialogo Sopra i due Massimi Sistemi del Mondo.

Correspondence with Fermat

When Pierre de Carcavy moved to Paris in 1636 and explained Fermat's ideas regarding Galileo's theory on the free fall of bodies to Mersenne, he immediately felt great curiosity for the mathematician's opinions and wrote him a letter. Fermat responded solicitously on 26 April of the same year. As well as explaining what he asked for, he informed him of his work on spirals, which arose from the study on the trajectory of the free fall of bodes, applying methods inspired by Archimedes' work *On Spirals*. He also spoke of his work restoring Apollonius' *Locis Planis*. He wrote:

> "I have also found many sorts of analyses for diverse problems, numerical as well as geometrical, for the solution of which Viète's analysis could not have sufficed. I will share all of this with you whenever you wish and do so without any ambition, from which I am more exempt and more distant than any man in the world."

A re-creation of the figure of Archimedes by Domenico Fetti, 1620. Fermat studied the work of this scientist from ancient Greece.

Fermat also took the opportunity to propose two problems on maximums to Marsenne, and asked him to pass the challenge of resolving them on to Parisian mathematicians. This first letter from Fermat to Mersenne demonstrates Fermat in his purest form.

Firstly, it is an example of his letter-borne relationship with the scientific community of his age, as letters were undoubtedly one of the essential media for the expression of his ideas. And secondly, he avoids any type of self-importance. Fermat served ideas and science, without giving import to his authorship. And finally his love for proposing new problems also appears, new mathematical challenges with which the abilities of his contemporaries could be measured. As he already had the answers – or so he said – the challenges that the mathematician set became doubly exciting.

When someone was able to resolve them it gave rise to the question of who deserved the credit for the first solution, but if nobody managed to solve them the value of the solution grew enormously as did the glory accruing to whoever eventually found it. And, of course, Mersenne was delighted to communicate the challenges set by Fermat to his salon members.

The cycloid problem

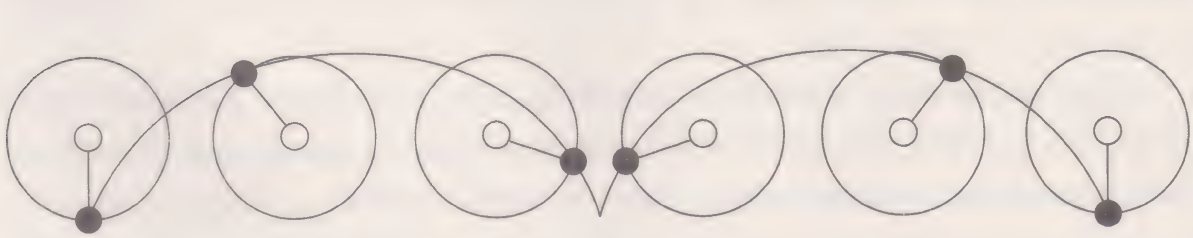
In 1632 Gilles de Roberval arrived in Paris to teach at the Collège Royale. Mersenne immediately appreciated his extraordinary talent for mathematics and got in contact to ask him some questions to which he still had not found the answer, among them was the cycloid problem. The chemistry of complicity was starting to take effect in resolving problems. In 1599 Galileo had defined the cycloid as a geometric shape described by the point of a circle as it rotates on a straight line.

Mersenne was immediately fascinated by the beauty of the curve and decided to study it. He was interested in some of its properties: the length of an arc of the curve, the area it covers, etc. In order to determine the area of the curve Galileo had carefully constructed a model from metal and weighed it on a set of scales. With this method he achieved a good approximation, but he was not satisfied, he wanted an exact mathematical answer.

Mathematical methods reside in the realm of ideas and can be said to be perfect. They are not limited by the imperfections of a metal model or imprecisions of a set of scales.

Mersenne studied the cycloid for years, publishing his results in various works: *Quaestiones in Genesim* (1623), *Mathematical Synopsis* (1626) and *Questions Inouyes* (1634). As usual he put the letter-writing system into action and communicated the results obtained and the problems he was able to resolve to his correspondents. Torricelli, Fermat, Descartes, Roberval, all correctly calculated that the area under the arc of the cycloid is the triple of ω by the square of the radius of the circle. Roberval and Wren calculated that the length of the arc equals eight times the radius. Answers with such elegance and simplicity, and with so many mathematicians behind them!

With this new way of working, implemented by Mersenne, all the talented minds were invited to resolve the interesting problems presented to them. The results were not the fruit of isolated thinkers, but of their collaboration with frequent exchanges of ideas. Although the story of science is often keen to attach names to things and



A diagram of the cycloid.

THE BRACHISTOCHRONE AND TAUTOCHRONE PROBLEMS

Let's suppose that we want to get from A to B in the quickest way possible, only subjected to the action of gravity. Or, in the same way, we want to find the shape of a curve that we are going to slide down to take us from A to B as quickly as possible. This curve is called brachistochrone (from the Greek *braquistos* – the shortest – and *cronos* – time).

At first, intuition tells us that the quickest path for sliding from A to B would be the shortest distance, in other words, a straight line. However, this is not the case. The fastest line is the inverted cycloid arc, which passes through A and whose minimum is at B.

In 1696 Johann Bernouilli already knew the solution to the problem and challenged that period's scientists to also find it. The problem was resolved independently by Leibniz, Newton, Jakob Bernoulli and L'Hospital.

In 1673 Huygens discovered that if you let an object fall on a cycloid arc, from whatever height, it will always take the same length of time to reach the base. Therefore, the cycloid also resolves the tautochrone problem (from the Greek *tauto* – the same – and *cronos* – time).

If the object is allowed to fall freely, it will take the same time to get from A to B as from A′ to B.

identify the people who first solved a problem, in this new working environment achievements often called for shared merit. Who could have imagined that such a beautiful curve that was once defined by Galileo from a circle rolling along a straight line would end up being proposed a few years later by Gérard Desargues for designing the teeth in the gears in engines, and would end up solving the brachistochrone and tautochrone problems!

The Ponte di Mezzo, in Pisa, designed by students of Galileo, with its three cycloid arches. The bridge was destroyed in 1944.

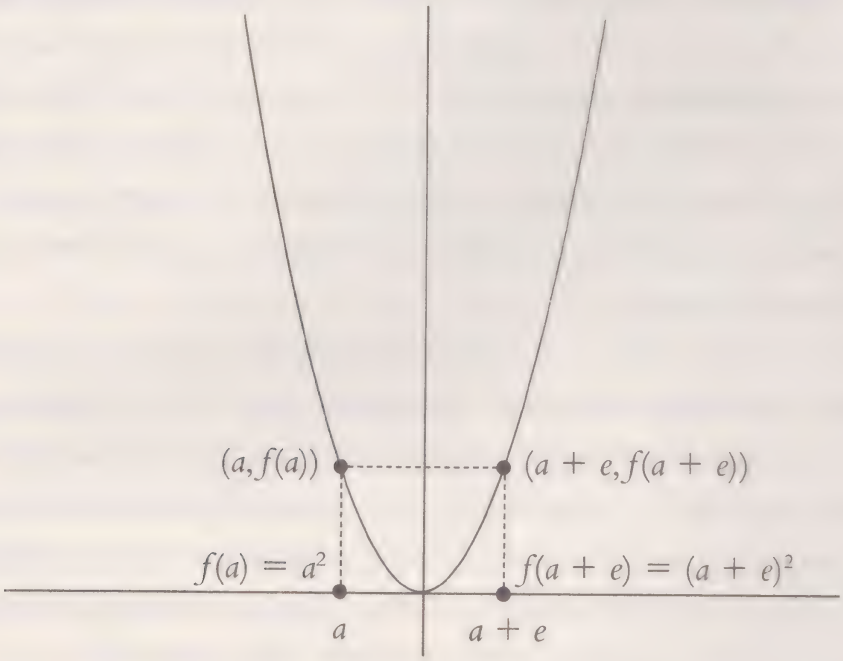
The maximums and minimums method

Roberval and Mersenne were intrigued by Fermat's results on maximums and minimums. The challenges set and resolved by Fermat in his letters could not be the result of coincidence, they had to be the result of mathematical methods that were unknown at that time. They realised that Fermat was far more advanced than his contemporaries in resolving problems on maximums and minimums, and they asked him to explain his methods. Fermat did not make them wait long for an answer and he sent them three texts (*Methodus ad Disquirendam Maximam et Minimam et de Tangentibus Linearum Curvarum*, *Apollonii Pergaei Libri Duo de Locis Planis Restituti* and *Ad Locos Planos et Solidos Isagoge*) for the consideration of the Parisian mathematicians. With it Fermat found fame at the forefront of mathematics.

In his *Maximums and minimums method* Fermat observed that when a function reached a maximum, then a straight line parallel to the x-axis cuts the function at one point only; and that in the values immediately below the maximum, a straight line parallel to the x-axis will cut the function at two very close points, one to the right and the other to the left of the maximum point.

Therefore, considering the extreme value of the function, $f(a)$, and a very close value, $f(a + e)$, where e is a very small amount, it can be deduced that these values must be practically equal, and therefore, according to Fermat, they can be 'adequalised'. An equation results from this process which, once the value of e is eliminated as it is negligible, allows the calculation of a.

Let's look at an example: the function $f(x) = x^2$. The graph for this function is shown in the illustration below.



A drawing of the parabola f(x) = x².

Suppose we want to calculate the minimum for this function $f(a) = a^2$. To do so we need to consider a very close value $f(a + e) = (a + e)^2 = a^2 + 2 \cdot a \cdot e + e^2$. Then they are 'adequaled', in other words they are considered to be so close that they are equal $a^2 = a^2 + 2 \cdot a \cdot e + e^2$. Next, the resulting equation is resolved. Eliminating a^2 from the two sides of the equation gives $2 \cdot a \cdot e + e^2 = 0$, and dividing by e gives $2 \cdot a \cdot e = 0$. Finally e is considered so close to 0 that it is negligible, leaving $2 \cdot a = 0$, and therefore $a = 0$, which correctly calculates the minimum of the given function.

As can be seen, after 'adequalling' we get an equation that is the equivalent of finding the derivative of the function and making it equal 0, in a time when it was not even known what deriving was or what it was to calculate limits of functions. So it is not surprising that some mathematicians had their reservations about the process of 'adequalling'. But Fermat's intuition on this subject is

completely brilliant, as if he were a magician of algebraic expressions who, in the end, spectacularly finds the desired result.

Multiplicity of interests

The restoration of Apollonius' lost works formed part of an ambitious mathematical project initiated by Viète and Marino Ghetaldi, a team to which Willebrord Snel and Fermat himself were later added. Fermat felt a deep admiration for the Greek mathematician. Fermat was also a great scholar of classical languages, and his broad knowledge of mathematics allowed him to take on the difficult task of translating Latin and Greek scientific works from their original authors with great aplomb. So it was not surprising that when a translator found a passage that was difficult to understand and was related to scientific or mathematical reasoning, they went to him in search of a solution.

In Toulouse he met Charles de Montchal, hellenist and great book lover, whose collection of manuscripts, with time, would form part of the Bibliothèque du Roi. Fermat had access to this exceptional archive. Montchal showed him *Les Harmoniques*, a work on music written by Byrene, a 14th-century Byzantine author. The mathematician read that text and wrote some annotations in the margins – as was his custom. While he read his imagination was flying, with many ideas occurring to him, and he noted them down in order to think about them later.

In Castres he met Pierre Saporta, who in 1664 wrote the *Traité Sur la Mesure des Eaux Courantes*, a French translation of a work by Benedictine father Benedetto Castelli. He also translated a work by Torricelli on the same subject into French, and in that translation he decided to include Fermat's description of an instrument for measuring the density of liquids described by Synesius, bishop of Cyrene, a contemporary of Hypatia of Alexandria. Castelli himself indicated that others before Fermat had tried to understand the workings of that strange instrument. Saporta admired Fermat's scholarship and, as well as dedicating the work to him, he had some glowing words for him: "The pages in this notebook remain empty and have given me the idea of filling them with an eloquent observation that I have learnt in the past few days from the incomparable M. Fermat, who does me the honour of holding me dear and often suffering my conversation." In another passage he writes: "All scholars of all types of literature consult you on difficult passages they find in books. I could put forward a large number of excellent observations that you have made on Synesius, Frontin, Athenée and many other authors,

and the clarifications which you have given on the obscure passages that could not be understood by the Scaligers, the Casaubons, the Petaus, and the Saummaises. Finally, it seems sir, that you were born to govern the empire of literature and to be the sovereign legislator of all scholars."

Fermat's erudition, therefore, was recognised by many of the people he knew. His interests encompassed all knowledge, with a profound encyclopedic style that did not discriminate between science and literature. Everyone had in Fermat a scholar who was prepared to listen and give his opinion. Thus, there were many people who approached him and asked for his collaboration.

Fermat wrote notes in many of the books that he read, leaving a record of his observations and thoughts. Everything fitted in the margins of the books. In fact, Fermat's literary legacy includes the annotations made in the margins of his own books. He even composed a poem with one hundred hexameter verses in Latin, *Cede Deo seu Christus moriens*, which he dedicated to Guez de Balzac and was read at the Académie de Castres in 1656.

In conclusion, Fermat had great respect and admiration for the classics, while he lived intensively in the search of the adventure of providing new ideas "that do not appear in books", neither ancient nor modern. He was a fervent reader of Francis Bacon, and was very enthusiastic about new scientific discovery. His universal spirit led him to recommend the union of all scientists, and he coincided with Mersenne in thinking that the more communication there is, and the more ideas contributed, the more science will advance.

Thus, for example, in 1657 he suggested to Cureau de la Chambre, regarding refraction, that "if you could allow me to unite a little of my geometry with your physics, we could do some novel work together."

Portrait of Francis Bacon. The English philosopher and proponent of scientific methods was greatly admired by Fermat.

Fermat also enjoyed the beauty of mathematics. He knew that behind the demonstrations and theories there was an underlying aesthetic and philosophy of knowledge. The way in which he presented his results and set his challenges indicates a person whose enthusiasm spread to his friends. On 25 December 1640, he wrote to Mersenne: "As soon as M. Frenicle writes to me, I will send you some proposals which I think will be of value to you, and with the utmost modesty, they are much more eloquent than anything we have talked about previously." Later, on 2 August 1641, Frenicle wrote to Fermat: "The methods that you provide [...] are truly very eloquent, and you have the means to make good use of your rules, which gives them a certain grace and makes them even more appreciated."

A strange way of working

Many historians ask themselves why Fermat noted so many things down but wrote so little. Why did he not write books to explain his ideas and discoveries? But the more we find out about Fermat the more we can understand that this was simply his way of working. He was not a professional mathematician. He liked to think about

mathematics, physics, literature, philosophy, music... and to make notes. It was as if he was having a conversation with the book he was reading and giving his opinion in the narrow margin which the page allowed. It was like thinking out loud. But he left the rigorous and complete demonstrations of all the possible cases to other people, who were perhaps more versed, or maybe had more time. To write a book would have needed dedication and time that he did not have. He dedicated his time to thinking and pondering new things. He had so many interests that he could not stop to write a book with all its introductions, basic explanations and details of the demonstrations. Once he had cleared up a subject, he noted it down and went onto another problem.

All this could give the image of Fermat as trivial, who dedicated time to solve one problem here and another there, both completely unrelated and with no basic purpose. His work did not seek to establish the grounds for a new mathematics in the style of Viète's *Introduction en L'art Analytic*, which seeks to provide a solution to all problems, or of Descartes' *Géometrie* which attempts to explain all phenomena of nature. But he was simply aware of the fact that his methods helped the advance of science and revolutionised working methods, providing new tools for resolving problems to which, previously, with old methods, no solution had been found.

Fermat's way of working is without doubt one of the ingredients that forms part of his legend. His way of solving problems brimmed with originality and creativity, but on occasions it is difficult to digest, and he did not satisfy all the mathematical details that some of his contemporaries asked of him. The type of writing that was best suited to his way

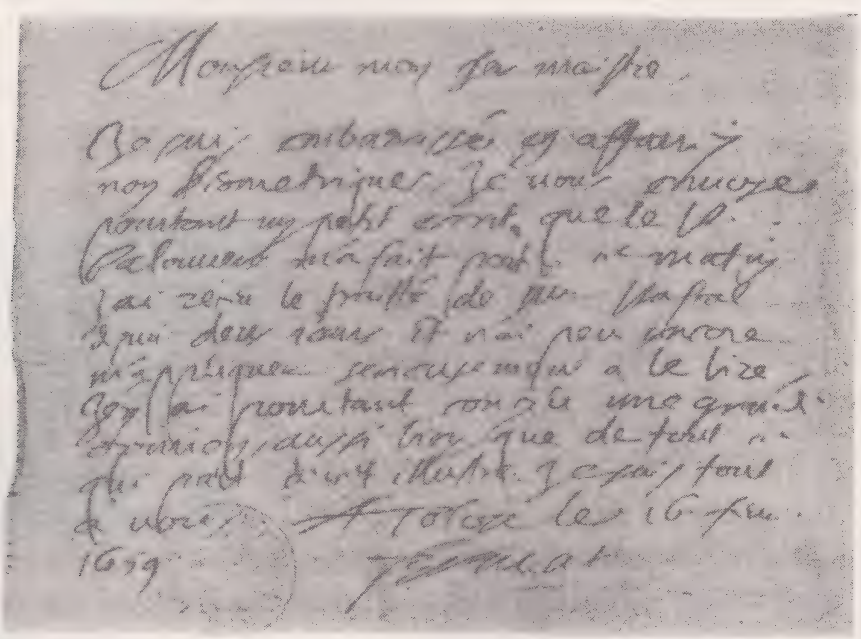

FRANCISCI VIETÆ
OPERA
MATHEMATICA,
In unum Volumen congesta,
ac recognita,
Operâ atque studio
FRANCISCI à SCHOOTEN Leydensis,
Matheseos Professoris.

LUGDUNI BATAVORUM,
Ex Officinâ Bonaventuræ & Abrahami Elzeviriorum.
cIↄ Iↄ c XLVI.

Frontispiece of one of the books by François Viète, who was well studied by Fermat, although both mathematicians had very different working methods.

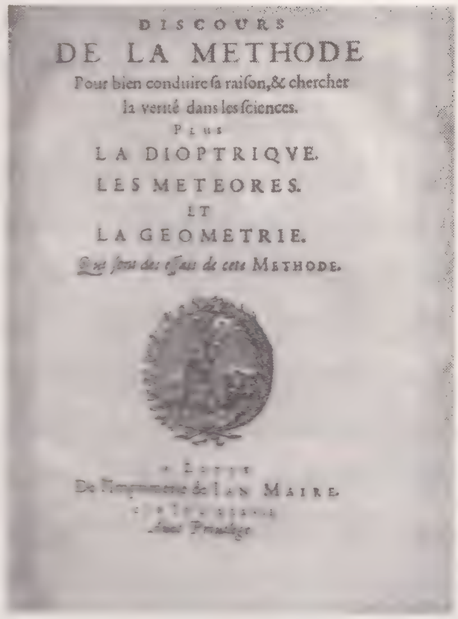
of thinking was letters. In them he could freely discuss science without necessarily following a formal discourse. They were also the ideal way to propose new challenges to his correspondents and to accept those proposed to him. If the occasion required it, he could go into more detail; although he normally only outlined his solutions, like clues so that the reader could continue to ponder the questions at hand alone, and as evidence that he actually had a solution, but he was not prepared to give it up easily. This resistance to explain his methods would form part of this exciting game of challenges, which culminated brilliantly with the challenge of his very last theorem.

Even so, the letters, despite their virtues, also had their downfall. Occasionally they would give rise to a misunderstanding, which would take years to resolve, or would simply never be resolved. Sometimes eternal arguments arose over whether a certain problem was first resolved by he who first said it or he who first wrote it. Occasionally the ideas were formed by many people and later many of them would claim sole authorship. Or sometimes a problem was resolved independently by several people and who deserved the accolade of being the first was discussed later. On top of all this some correspondents found fruit in all these issues, or sometimes there was a 'slip of the pen' that had disastrous consequences for anyone who found themselves involved. Fermat undoubtedly tried to stay out of these issues but, despite his efforts, he was not unaffected by them.



Hand-written letter from Fermat.

The dispute with Descartes



Cover of Descartes' famous Discourse on Method *which includes the paper on dioptry that led to the clash with Fermat.*

At the beginning of 1637, through Mersenne René Descartes requested permission from the king of France to publish his *Discours de la Méthode* and his three *Essais*. Beaugrand then occupied the position of secretary to the chancellor and, therefore, he was in a position of responsibility that gave him influence over the matter. The chancellor, Pierre Seguier, retained a copy of Descartes' first draft of *La Dioptrique*, in order to have it evaluated. Beaugrand, without the author's permission and without informing Mersenne, sent a copy to his friend Fermat so that he could give his opinion on it. Mersenne found out later and wrote to Fermat asking him to be discrete and that any comment he would like to make on the matter be communicated directly to him and no one else. On 22 September 1637, Fermat wrote to Mersenne: "You asked for my opinion on M. Descartes' *La Dioptrique* (*Optics*). This is my impression of these new propositions, although the consequences that he deduces regarding the shape of the lenses should have are eloquent, it would be desirable that the fundamentals on which they are established were better tested; but I fear the truth just as absent as the test." In the same letter Fermat gave the reasons on which he based his opinion. In *La Dioptrique* a model is proposed for the nature of light to explain the law of refraction. But this model makes assumptions that in Fermat's opinion were not sufficiently justified. Fermat pointed out a contradiction between the idea of the instant propagation (or transmission) of light and that the speed of light depended on the medium in which it was propagated; nor was it clear to him that light travelled faster as the density of the medium increased.

Mersenne passed Fermat's opinions on to Descartes, and it can be assumed that Descartes did not like them. So, in October 1637 he responded to Mersenne: "The error which M. Fermat found in my demonstration (on refraction) is no more than imaginary and demonstrates well that he has done no more than browse through my treatise."

But Fermat was not looking for an argument, just to discover the truth. In a letter written to Mersenne in December 1637 he proposed a new model for refraction. In that letter he wrote: "[…] it is not for envy or rivalry that I continue with this small dispute, but just so that the truth may come to light; which I am sure will please M. Descartes, whose merit is so well known and which I expressly recognise here."

But the dispute had already begun. For Descartes, Fermat's opinions were a challenge to his status and his ideas. His *Discours de la Méthode* and his three *Essais* were the basis of his philosophy and the grounds for his thinking. So he decided to prepare himself for fully fledged battle. On 18 January 1638, he wrote to Mersenne: "If the author is surprised that I do not have such rules in my geometry, I have far more reason to be surprised that he wants to fight with mean guns. But I would like to give him more time to get back on his horse and use all the improvements that he has chosen for this combat."

When Fermat sent his *Methodus ad Disquirendam Maximam et Minimam et de Tangentibus Linearum Curvarum* for the consideration of the Parisian mathematicians, Descartes saw the opportunity to get even and he accused Fermat of doubtful reasoning. Roberval and Étienne Pascal came out in defence of Fermat, while at first Mydorge and Desargues took Descartes' side. In April 1638 Roberval wrote: "When M. Descartes has properly understood M. Fermat's Maximums and Minimums Method, then he will no longer be surprised that this method is defended by some people, and he will admire the method in itself, which is excellent and worthy of its author." Mersenne's role in all of this is interesting, because in reality all of this came through him. Both Descartes and Fermat sent letters to Mersenne, taking for granted that he would then explain them to the other party. Finally, Desargues agreed with Fermat, and Descartes ended up accepting the verdict of the evidence and recognising that "seeing the latest method he uses for finding the tangents of curves, I cannot respond in any other way but to say that it is very good, and that if he had explained it in this way at the beginning there would have been no contradiction whatsoever."

Little by little things began to calm down. On 29 June 1638, Descartes wrote to Mersenne: "I have seen that you were kind enough to inform me of the letters

that M. Fermat has written to you; and firstly regarding that which he said about finding words in my first article which were more bitter than expected, I humbly ask for his forgiveness, and that he consider that I did not know him." Finally, in October 1638, Descartes wrote, for the first time, a letter addressed to Fermat, as an apology: "I must frankly confess here that I have never known anyone who has given me the impression of knowing as much about geometry as you... However, we pay more attention to small imperfections in diamonds, in contrast to great blemishes in ordinary stones, thus I think I should look more closely at things that come from you than that which comes from another person who I think less of."

However, the incident did not end there. Despite his words, Descartes saw Fermat as a genius and a rival, and therefore he feared him and tried to discredit him at any opportunity. On one occasion, having analysed Fermat's work on the determination of a tangent of a cycloid, which was in fact correct, Descartes wrote to Mersenne saying that his work had errors, and that Fermat was an inadequate mathematician and thinker. Descartes' influence within the scientific community of his times meant that many scientists of the time had a distorted image of Fermat.

But Fermat's genius just kept on shining through. In fact, he was the first to establish the bases of algebraic geometry, before Descartes published his *Géometrie*. Together with Pascal he founded the theory of probability. In arithmetic his results and methods of demonstration marked the beginning of the modern theory of numbers. And these are just a few examples. In short, as a mathematician there was no doubt that Fermat eclipsed Descartes. Even so Fermat attempted to cool the situation, and with a fine sense of humour, regarding an error he had seen in *La Géometrie*, he said that such was his admiration for Descartes' genius that, even when he made mistakes, his work was worth more than that of others with no mistakes.

The theory of refraction

In terms of the theory of refraction, the story went on. After Descartes' death, one of his students wanted to prepare a publication with all his correspondence. So he asked for Fermat's collaboration in order to gather all the letters that both had written. This motivated Fermat to go back and revise his work on refraction, and not happy with his own arguments he investigated the subject again. It was then that he established the principle in which light follows minimum-time trajectories. Known as the 'Fermat principle', it was published in *Analyse Pour les Réfractions et Synthèse pour les Réfractions*, which dates from approximately 1660, and from it he

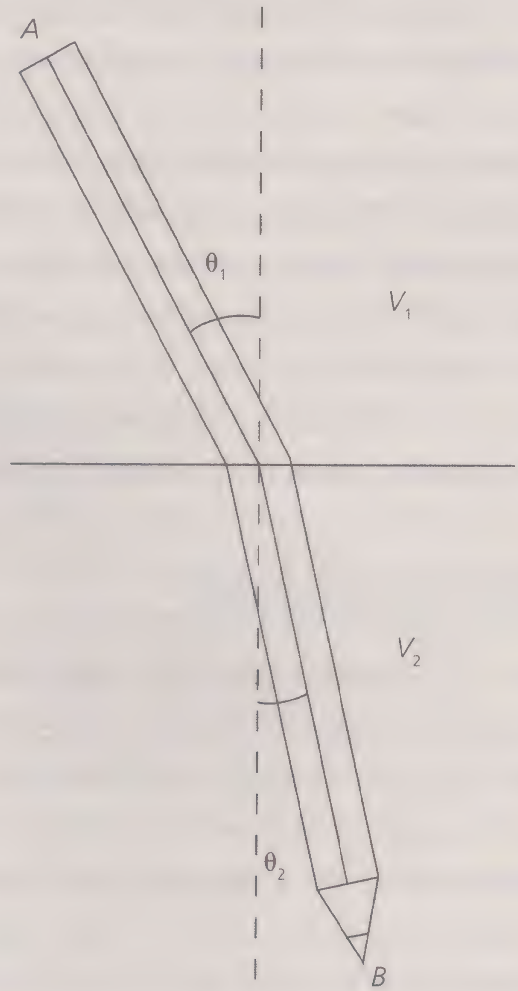
gives a mathematical explanation of Snell's Law. It is another example of Fermat's perseverence in resolving problems he set himself. He did not mind returning to them over and over again in order to add to them. This perseverance is the same that his contemporaries expected in resolving the challenges set for each other.

SNELL'S LAW

When a stick is put into water it gives the impression of being broken or bent. Its angle under the water seems different to the part above the water. This optical effect, called refraction, is due to the fact that the velocity of light changes depending on the density of the material through which it is travelling. Air is less dense than water, and the speed of light is greater in the air than in water as it has fewer obstacles in its path.

Willebrord Snell found a formula, known as Snell's Law, which related the speed of light in the two media and angles of incidence and refraction:

$$\sin \theta_1 / v_1 = \sin \theta_2 / v_2$$



Fermat's principle provides an explanation for this phenomenon. According to this principle, light follows minimum-time trajectories. Suppose, as shown in the diagram, that a bird wants to go from point *A* at the end of the stick that is out of the water to point *B* at the end of the stick that is submerged.

Suppose that in air it flies at a velocity v_1, and that in the water it dives at a different speed, v_2. Fermat showed that the fastest way of going from *A* to *B* was not a straight line as described by the stick leaning at an angle. Therefore, the bird will have to change direction in order to get to *B* as fast as possible.

Chapter 4

The Birth of the Last Theorem

It is impossible for a cube to be the sum of two cubes, a fourth power to be the sum of two fourth powers, or in general for any number that is a power greater than the second to be the sum of two like powers. I have discovered a truly marvellous demonstration of this proposition that this margin is too narrow to contain.

Pierre de Fermat

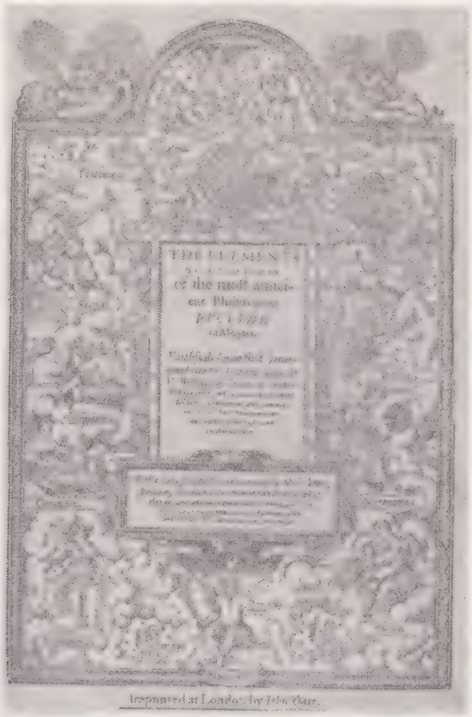
One day a copy of Diophantus' *Arithmetica* fell into the hands of Fermat. While Fermat read the work, his mind traversed landscapes of extraordinary mathematical beauty, while he came up with some more of the most devilish problems that he set for the mathematical community. Among them is the famous Last Theorem. Of all Fermat's challenges, it is the one that has taken by far the longest to solve. Fermat wrote it in the margin of the page of the second volume containing problem 8, as we will see in detail below.

The roots of Fermat's enigma on the theorem of numbers can be traced back to the Alexandrian era. On the one hand, Euclid's *Elements* which dates back to the 2nd century B.C., and on the other, Diophantus aforementioned *Arithmetica*, which is from around five or six centuries later. Practically all the mathematical research in the Mediterrenean and Middle East for around 1,500 years was built up on the basis of these two books.

Euclid's *Elements*

Euclid's *Elements of Geometry* includes three books on arithmetic. These three books (volumes 5, 6 and 7) contain the first general theory on divisibility. It already mentions the maximum common denominator and an algorithm for calculating it, which is known by the name of 'Euclid's algorithm'. It also defines the prime numbers, and

demonstrates that they are infinite; it talks of co-prime numbers, and defines perfect numbers as those that are obtained as a sum of their own divisors.



Front cover of the first English version of Euclid's Elements *dating from the year 1570.*

PYTHAGOREAN TRIPLES

In Euclid's *Elements* a general formula is given for obtaining Pythagorean triples, in other words, natural numbers that satisfy Pythagoras' equation $a^2 + b^2 = c^2$. In it, natural numbers m, n, where m is greater than n are arbitrarily selected and then the following calculation is made:

$$a = m^2 - n^2 \,;\, b = 2mn \,;\, c = m^2 + n^2.$$

Numbers a, b, c which are obtained verify that:

$$a^2 + b^2 = (m^2 - n^2)^2 + (2mn)^2 = m^4 - 2m^2n^2 + n^4 + 4m^2n^2 = m^4 + 2m^2n^2 + n^4 = (m^2 + n^2)^2 = c^2,$$

and therefore they form a Pythagorean triple. If we also take m, n where they are coprime and only one of them is even, then the formula generates all the primitive Pythagorean triples, in other words, those in which a, b, c are coprime. From there it is deduced that there are an infinite number of primitive Pythagorean triples.

Each Pythagorean triple allows us to draw a right-angle triangle with integer sides. Fermat demonstrated that the area of these triangles cannot be a square number.

Perfect numbers

The history of perfect numbers could command its own chapter – the story of mathematics. The search for these numbers can be compared in a certain sense with the search for the decimals of π. Initially a few were found, and as mathematical knowledge advanced, more and more were revealed. But many mistakes were made on the way, which were corrected with time. Nowadays, in the modern era of computers, they have still not all been found and, in fact, it is not known if the total of these numbers is finite or infinite.

The term 'perfect' alludes to an aesthetic concept of mathematics. The beauty of these numbers does not reside in the way they look, in the difficulty of finding them, or in convoluted definitions, but in one simple property.

Take the number 6 for example. Its divisors, that is to say, the numbers that can divide into it exactly, are 1, 2, 3 and 6. Marvellously $1 + 2 + 3 = 6$, in other words, adding all the divisors save the largest gives the result of 6. The next perfect number is 28; its divisors are 1, 2, 4, 7, 14 and 28, and $1 + 2 + 4 + 7 + 14 = 28$. The next perfect number is 496; its divisors are 1, 2, 4, 8, 16, 31, 62, 124, 248 and 496, and $1 + 2 + 4 ++ 8 + 16 + 31 + 62 + 124 + 248 = 496$. The next is 8,128, as $1 + + 2 + 4 + 8 + 16 + 32 + 64 + 127 + 254 + 508 + 1{,}016 + 2{,}032 + 4{,}064 = = 8{,}128$. These four perfect numbers have been known since ancient times; Euclid includes them in his *Elements*, and in proposition 36 of Book IX gives a formula for generating them.

The generation of perfect numbers

In the year 100 B.C., neo-Pythagorean philosopher Nicomachus wrote his *Introductio Arithmetica*, in which he classified numbers as abundant (those where the sum of its own divisors is greater than the number itself), deficient (those where the sum of its own divisors is smaller than the number itself) and perfect (those where the sum of its own divisors exactly equal the number itself). In his treatise he explained the formula given by Euclid to find them, "which does not leave out any perfect numbers and which does not include any that are not; and which is done in the following way. First set out in order the powers of two in a line, starting from one, and proceeding as far as you wish: 1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048, 4096; and then they must be totalled each time there is a new term, and at each totalling examine the result, if you find that it is prime and non-composite, you must multiply it by the quantity of the last term that you added to the line, and the product will always be perfect. If,

otherwise, it is composite and not prime, do not multiply it, but add on the next term, and again examine the result, and if it is composite leave it aside, without multiplying it, and add on the next term. If, on the other hand, it is prime, and non-composite, you must multiply it by the last term taken for its composition, and the number that results will be perfect, and so on as far as infinity." In other words, or rather, in numbers, we can say:

$1 + 2 = 3$ is a prime number and therefore $(1 + 2) \cdot 2 = 3 \cdot 2 = 6$ is perfect,
$1 + 2 + 4 = 7$ is a prime number and therefore $(1 + 2 + 4) \cdot 4 = 7 \cdot 4 = 28$ is perfect,
$1 + 2 + 4 + 8 = 15$ is not a prime number and therefore we skip it, and then
$1 + 2 + 4 + 8 + 16 = 31$ is a prime number, and therefore,
$(1 + 2 + 4 + 8 + 16) \cdot 16 = 31 \cdot 16 = 496$ is perfect,
$1 + 2 + 4 + 8 + 16 + 32 = 63$ is not a prime number and therefore we skip it, and finally
$1 + 2 + 4 + 8 + 16 + 32 + 64 = 127$ is a prime number, and therefore
$(1 + 2 + 4 + 8 + 16 + 32 + 64) \cdot 64 = 127 \cdot 64 = 8{,}128$ is perfect.

Of course, this formula correctly finds the first four perfect numbers. There is another, simpler formula that also generates them. It is easy to see that when the powers of two, starting with 1 and without skipping any, are added, the result obtained is the following power of two minus one. In numbers:

$$1 + 2 = 3 = 4 - 1 = 2^2 - 1;$$
$$1 + 2 + 4 = 7 = 8 - 1 = 2^3 - 1;$$
$$1 + 2 + 4 + 8 = 15 = 16 - 1 = 2^4 - 1.$$

And so on. In this way, with current notation, Euclid's formula for calculating perfect numbers can be translated:

$$6 = (2^2 - 1) \cdot 2$$
$$28 = (2^3 - 1) \cdot 2^2$$
$$496 = (2^5 - 1) \cdot 2^4$$
$$8{,}128 = (2^7 - 1) \cdot 2^6.$$

And as long as $2^n - 1$ is a prime number $(2^n - 1) \cdot 2^{n-1}$ will be a perfect number.

Conjectures regarding perfect numbers

In ancient times, mathematician used the first four known perfect numbers to venture a series of conjectures. For example, it can be seen that the value of n for the first four perfect numbers follows the succession of the prime numbers 2, 3, 5, 7. It was tempting to think that the next perfect number would be $(2^{11} - 1) \cdot 2^{10}$, but actually it was not, because $2^{11} - 1 = 2{,}047 = 23 \cdot 89$ is not a prime number and therefore, $n = 11$ does not generate a perfect number.

It was also observed that the first perfect number had one digit, the second had two, the third three, and so on. Therefore, it was thought that the fifth perfect number should have five digits. But this is not the case as the fifth perfect number turned out to be $(2^{13} - 1) \cdot 2^{12} = 8{,}191 \cdot 4{,}096 = 33{,}550{,}336$, which has 8 digits.

Another observation made by mathematicians in ancient times was that the last digit of the first four perfect numbers alternated between the numbers 6 and 8. So the sixth perfect number should finish in 8. But this was not true either, because the sixth perfect number turned out to be $(2^{17} - 1) \cdot 2^{16} = 131{,}071 \cdot 65{,}536 = 8{,}589{,}869{,}056$, which ends in 6.

But not all of these conjectures turned out to be false. It was also inferred that all perfect numbers were even, and that the given formula could provide all of them. Incredibly easy to say, but incredibly difficult to demonstrate. It was not until the 18th century and the work of Leonhard Euler that it was first demonstrated that all the even perfect numbers could be found in this way; therefore, it was demonstrated that, although not alternating, all even perfect numbers finish in six or in eight. But even today we still do not know if there are odd perfect numbers, The most that has been demonstrated is that, if there are any odd perfect numbers they must be greater than 10^{300}. However, that does not mean that there are absolutely no odd perfect numbers,

Portrait of Leonhard Euler by Emanuel Handmann. This 18th century mathematician made important discoveries in the field of perfect numbers and prime numbers.

because, what is the significance of a few centillion centillion centillions compared to the immense infinity of natural numbers?

It was also inferred that there were an infinity of perfect numbers, but this has not been demonstrated either. Every so often, the discovery of new Mersenne prime numbers is announced, and every Mersenne prime which is discovered leads to the

FERMAT'S PRIME NUMBERS AND SUBSEQUENT ADVANCES

Portrait of Carl Friedrich Gauss.

In 1650 Fermat set the mathematical community one of the most famous challenges in history when he stated that all numbers of the type $F_n = 2^{2^n} + 1$ were prime numbers. Everything actually seemed to indicate that he was right. $n = 0$ gave $F_0 = 3$, which is a prime number. $n = 1$ gave $F_1 = 5$, which is also a prime number. And also $F_2 = 17$, $F_3 = 257$ and $F_4 = 65,537$ are prime numbers. No more advances were made until 1732 when Euler demonstrated that $F_5 = 4,294,967,297 = 641 \cdot 6,700,417$, and, therefore, is not a prime number. It then took until 1880 for Landry to find a factorisation for $F_6 = 274,177 \cdot 67,280,421,310,721$, a real feat in a time when all calculations were done manually. In 1975 Morrison and Brillhart took a step forward by factorising $F_7 = 59,649,589,127,497,217 \cdot 5,704,689,200,685,129,054,721$, using computers. To date no more of Fermat's primes have been found, and neither has anyone been able to prove that there are no more; but calculating the prime factors that factorise them is a daunting task.

Why are we interested in whether numbers of this form are prime numbers or not? One answer was provided by Gauss when he demonstrated that for a regular polygon to be drawn in a circle using only a ruler and compasses, the number of sides must be such that its factorisation only contains twos and different Fermat numbers.

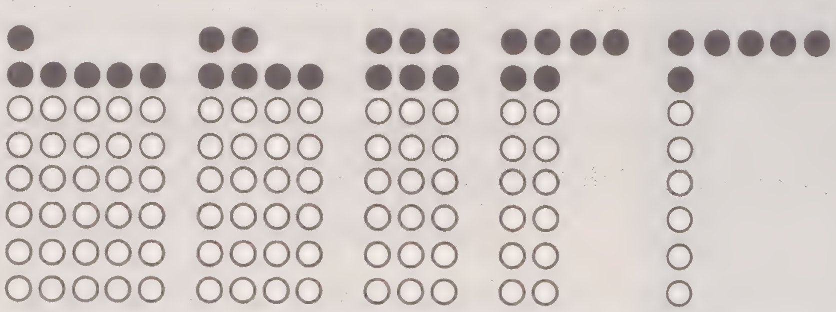
So, for example, with a ruler and compasses a triangle (three sides), a square ($4 = 2^2$ sides), a pentagon (5 sides), a hexagon ($6 = 2 \cdot 3$ sides), an octagon ($8 = 2^3$ sides) or a decagon ($10 = 2 \cdot 5$ sides) can be drawn, but not a heptagon (seven sides is not a Fermat prime number) or a nonagon ($9 = 3^2$ sides is the product of two equal Fermat prime numbers). Although there are procedures for approximate drawings for these cases, there will never be any exact ones.

corresponding perfect number. Hundreds of volunteers currently participate in the GIMPS (Great Internet Mersenne Prime Search) project, aimed at finding Mersenne prime numbers, offering their computers to run a program thought up by George Woltman. The result of all this joint work led to the announcement in August 2008 that the largest Mersenne prime known to date was, $2^{43112609} - 1$, which led to the

THE ARAB PEARL PROBLEM

Malba Tahan, the pseudonym of Júlio César de Mello e Souza, describes a beautiful problem in his book *The Man who Counted*, published in 1949. "A rajah willed his daughters a certain number of pearls and ordered that they be distributed as follows: his first daughter would receive one pearl plus one seventh of those remaining, his second daughter two pearls plus one seventh of those remaining, his third daughter would get three plus one seventh of the remainder. And so forth for the rest of his daughters. The youngest complained to a judge, alleging that the complicated system was greatly unfair to her. The judge, who, as tradition has it, who was skilled in resolving problems, quickly responded that the prosecutors were wrong and that the division proposed by the rajah was fair and perfect. The judge was right. Once the division was made, each of the sisters received the same number of pearls." The question is, how many pearls were there and how many daughters did the rajah have?

The solution is very simple: there were 36 pearls and they had to be divided between 6 people. The first received one pearl plus one seventh of 35, or 5; so 6 pearls were handed out, leaving 30. The second, from the 30 she found took 2 plus one seventh of 28, which is 4; so she took 6 and left 24. The third, from the 24 she found, took 3 plus one seventh of 21, which is 3 and, therefore, she left 18. The fourth, from the 18 she found, took 4 plus one seventh of 14, or 2; she also received 6 pearls. The fifth found 12 pearls; from those 12 she took five 5 plus one seventh of 7, or 1; so she took 6. Finally, the youngest daughter was left with the remaining 6 pearls. When setting enigmas Arab sensibilities often combine a literary beauty with a mathematical one. An inheritance of 36 marvellous pearls for six beautiful daughters, because six is a perfect number and thirty six is the square of a perfect number.



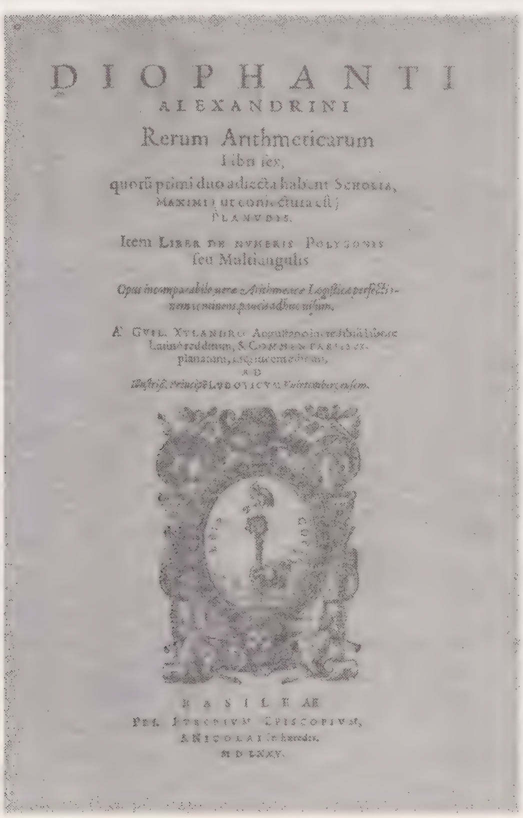
Graphical representation of the Arab pearl problem. (Source: The Man who Counted, *by Malba Tahan.)*

largest known perfect number, $2^{43112608} \cdot (2^{43112609} - 1)$, a number with 25,956,376 digits! On 12 June, 2009, it was announced that a new Mersenne prime had been found which was a tiny bit smaller, $2^{42643801} - 1$, which led to the forty-seventh perfect number discovered, $2^{42643800} \cdot (2^{42643801} - 1)$, which only has 25,674,128 digits! And although they get more and more scarce, and you have to dig further to find them, no one knows if this set of numbers is actually infinite. The GIMPS project continues to look for them.

Diophantus' *Arithmetica*

Practically nothing is known about the life of Diophantus – it is not even very well known when he lived exactly. However, we do have some clues. Firstly, Diophantus refers to the work of Hypsicles when he gives the definition of polygonal numbers, and therefore, his work was written after the year 150 B.C. Also, Theon of Alexandra, the father of Hypatia, quotes at least one of Diophantus' definitions in his work, from which we can deduce that he wrote *Arithmetica* before 350 A.D. Thus, for the moment, there is still a 500-year margin of doubt over Diophantus' dates of birth and death.

A better clue about the life and times of Diophantus is a letter from Michael Psellus, Byzantine author from the 11th century. The translation from Greek proposed by Thomas Heath is as follows: "Diophantus dealt with it [Egyptian arithmetic] more accurately, but the very learned Anatolius collected the most essential parts of the doctrine as stated by Diophantus in a different way and in the most succinct form, dedicating his work to Diophantus." Paul Tannery published this letter in one of his investigations and suggested that Psellus was referring to a comment about Diophantus whose original was lost and that was possibly written by Hypatia. Given that the Ana-



Cover of the version of Diophantus' Arithmetica *printed in Basel in 1575.*

tolius referred to in the letter is the Bishop of Laodicea, a writer and teacher of mathematics who lived in the 3rd century, it was deduced that Diophantus wrote his *Arithmetica* in around 250 A.D. However, not all researchers agree with this translation, because the initial date cannot be considered to be definite.

As in the case of Fermat, there is an epitaph that specifies Diophantus' age. It comes from the *Greek Anthology*, compiled by Metrodorus in around 500 A.D. This collection of riddles contains one about the author of *Arithmetica* that says:

> "His childhood lasted 1/6 of his life; he was married after another 1/7; his beard grew after another 1/12 and his son was born 5 years later; the son lived half the years of the father, and the father died four years after the son."

If we make Diophantus' age x this gives that his childhood lasted $x/6$ years; he married after $x/7$ years; his beard grew for $x/12$ years; his son was born five years later; his son lived $x/2$ years; and the father died 4 years later. Which gives:

$$x = x/6 + x/7 + x/12 + 5 + x/2 + 4.$$

Multiplying both sides of the equation by 84 gives:

$$84 \cdot x = 84 \cdot x/6 + 84 \cdot x/7 + 84 \cdot x/12 + 84 \cdot 5 + 84 \cdot x/2 + 84 \cdot 4.$$

Simplifying gives:

$$84 \cdot x = 14 \cdot x + 12 \cdot x + 7 \cdot x + 420 + 42 \cdot x + 336.$$

And grouping the x s together gives:

$$84 \cdot x - 14 \cdot x - 12 \cdot x - 7 \cdot x - 42 \cdot x = 420 + 336.$$

Which gives $9 \cdot x = 776$ and, therefore, $x = 756/9 = 84$. Thus, Diophantus married at 26 and had a son at 38, who lived 42 years, half the age which he lived. However, it is unknown if this problem is purely an invention or if it is actually based on the mathematician's life.

Importance of the work

It is difficult to contemplate the importance of Diophantus' work. The problems he proposed are a challenge to ingenuity and creativity, and are a celebration of the beautiful aesthetics of mathematics. Although Diophantus did not use sophisticated

THE BOOKS OF DIOPHANTUS' *ARITHMETICA*

Diophantus' *Arithmetica* consists of thirteen books, of which six remain, all in the original Greek. In 1972 an Arabic manuscript appeared containing four more books, from an Arab source, which do not coincide with any of the Greek books. In them a long series of problems are described in which rational solutions are sought for polynomial equations for rational coefficients. The books from the Greek source cover a total of 189 problems. They are distributed as follows:

Book I: Contains 25 linear problems and 14 quadratic ones.

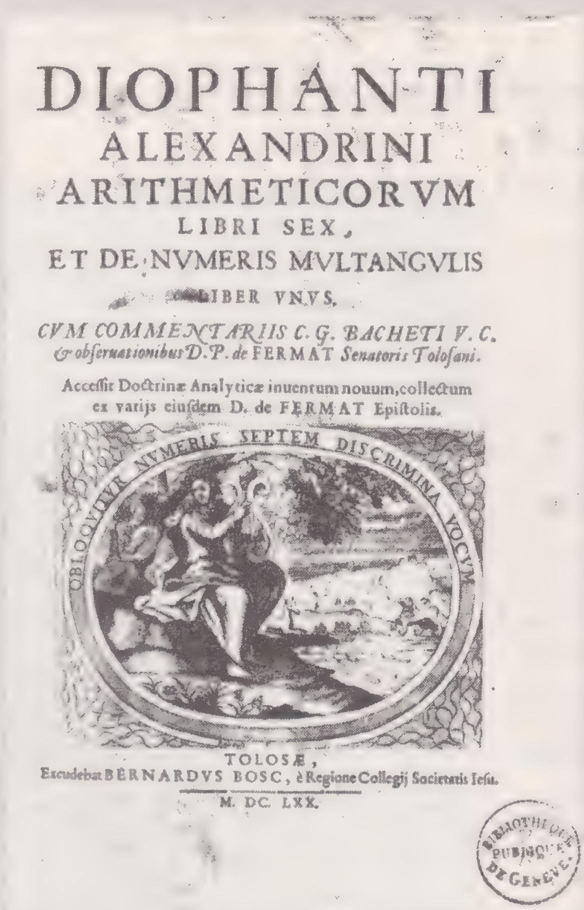
Book II: Contains 35 problems. Problem 8, undoubtedly the most famous, gave rise to the so-called 'Fermat's Last Theorem'.

Book II: Contains 21 problems. The most famous is number 19, in which geometry is used to find the solution for the first time.

Book IV: Contains 40 problems, most of which refer to cubic numbers.

Book V: Contains 30 problems. Most of them (28) are quadratic and cubic problems. In the last, number 30, Diophantus proposes a mixed problem.

Book VI: Contains 24 problems. It is dedicated to solving right-angle triangles with rational sides.


DIOPHANTI
ALEXANDRINI
ARITHMETICORVM
LIBRI SEX,
ET DE NVMERIS MVLTANGVLIS
LIBER VNVS.
CVM COMMENTARIIS C. G. BACHETI V. C.
& obseruationibus D. P. de FERMAT Senatoris Tolosani.
Accessit Doctrinæ Analyticæ inuentum nouum, collectum
ex varijs eiusdem D. de FERMAT Epistolis.

TOLOSÆ,
Excudebat BERNARDVS BOSC, è Regione Collegij Societatis Iesu.
M. DC. LXX.

Front cover of an edition of Diophantus' Arithmetica *which was published in 1670 by Fermat's son after the death of his father and includes the mathematician's famous comments.*

algebraic notation, he did introduce an algebraic symbolism, using an abbreviation for the unknown and powers of the unknown. This would allow him to express the equations more easily. He also used an abbreviation for the word *equals*. Therefore, his work symbolises a definitive step from verbal algebra to symbolic algebra.

From the problems dealt with in *Arithmetica* it has also been learnt that Diophantus was more involved in particular cases than general ones. Evidently, the advance that this already implicated was a giant leap, but also a few of his specific methods for resolution can easily be extended to more general cases. However, it is obvious that he lacked the most powerful algebraic notation for expressing more general methods. For example, Diophantus only had notation for one unknown, and when a problem implicated the appearance of several unknowns the name they were given was 'first unknown', 'second unknown', 'third unknown'... Nor did he have a symbol for expressing any number n so the expression $(6n + 1) / (n^2 + n)$, is expressed in words: "A sixfold number increased by one, which is divided by the difference between the square and the number itself." It is easy to see how it was

VIÈTE'S ALGEBRAIC NOTATION

Nowadays mathematics without symbolic notation is unthinkable. But the development of this notation system is the result of thousands of years of effort. Diophantus and Euclid used letters to denote their demonstrations. But it was Viète who made a definitive leap forward in algebraic notation. In his work *In Artem Analyticem Isagoge (Introduction to the Analytical Art)*, from 1591, Viète highlighted the interest in algebraic methods and tried to create a systematic demonstration. His proposal was to contrast the synthetic method for demonstrating theories used by the Greeks with a new method for discovering how to resolve problems. This created a new focus on what was already known in an attempt to leave no mathematical problem unsolved. Viète had no doubt in saying that thanks to algebra all problems would be solvable.

The evolution of the annotation system can be seen in the following example of a polynomial written with Diophatus' abridged notation, Viète's symbolic notation, and modern notation as it stands today.

Diophantus' abridged notation: $\mathrm{K}^{\Upsilon}\mathrm{K}\bar{\alpha}\Delta^{\Upsilon}\Delta\bar{\alpha}\Delta^{\Upsilon}\bar{\alpha}\mathrm{M}\alpha \uparrow \mathrm{K}^{\Upsilon}\Delta\bar{\alpha}\mathrm{K}^{\Upsilon}\bar{\alpha}\varsigma\bar{\alpha}$.

Viète's symbolic notation: $CC - CQ + QQ - C + Q - N + 1$.

Modern notation: $x^6 - x^5 + x^4 - x^3 + x^2 - x + 1$.

difficult to develop complicated expressions in this way. Mathematics would have to wait for Viète to make the final step from Diophantus' abridged notation and modern algebraic notation.

Diffusion of Diophantus' legacy

Diophantus' mathematical legacy was revealed to European mathematicians by German astronomer Johann Müller, also known as Regiomontanus, who in around 1463 discovered a copy of *Arithmetica* in Venice and noted that "nobody had yet translated Diophantus' thirteen books from Greek to Latin, in which the true flower of all mathematics lies hidden." In around 1570, Rafael Bombelli translated part of *Arithmetica*, although it was never published. However, he borrowed many of Diophantus' problems for his work entitled *Algebra*. In 1575 Wilhelm Holzmann, also known as Guilielmus Xylander, published *Diophanti Alexandrini Rerum Libri Sex* in Basel – it was the first Latin translation of Diophantus' work. In 1621 Bachet de Méziriac went one step further and published a new translation in Paris under the title: *Diophanti Alexandrini Arithmeticorum Libri Sex, et de Numeris Multangulis Liber Unus. Nunc Primun Graece et Latini Editi atque Absolutissimis Commentariis Illustrati.* This

A portrait of Johann Müller, who discovered a copy of Diaphantus' work in the 15th century.

HYPATIA OF ALEXANDRIA

Hypatia's life is surrounded by legend. Her date of birth is highly disputed and although it is known that she died in the year 415 A.D., historians cannot agree on her age at the time of her death. Theon, her father, was a well-known scholar and professor of mathematics from Alexandria, and he brought Hypatia up surrounded by his enthusiasm for the sciences. He also taught her about the different religions of the world and showed her how to follow a physical routine for keeping her body healthy and strong. Very quickly, Hypatia became an excellent speaker, and many people visited from other cities to study with her. Her students were a group of close-knit pagan aristocrats and Christians, some of whom held senior positions. The philosopher Dimascius said that "as well as achieving a higher level of practical virtue in the art of teaching, she was fair and wise, and was a virgin for her whole life."

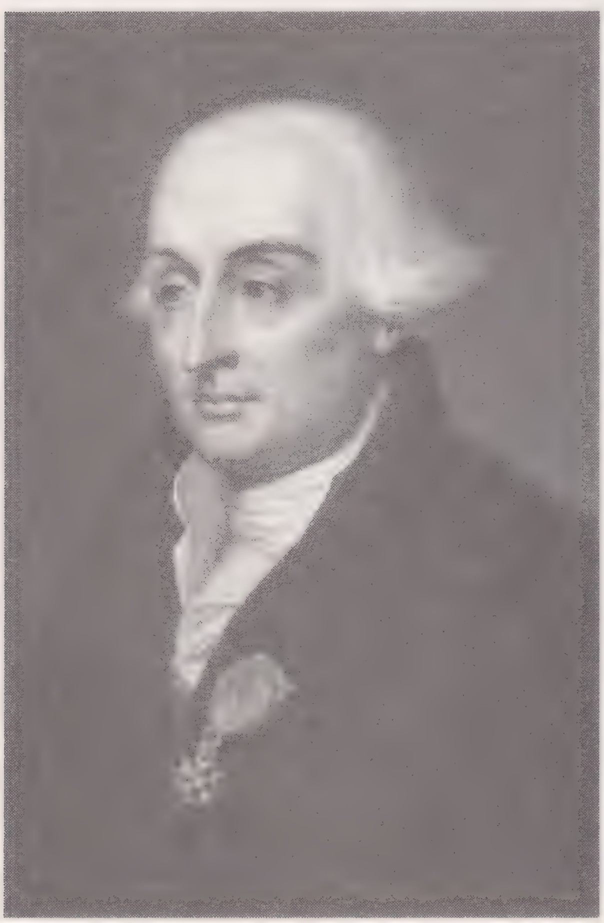
Hypatia's studies included astronomy, astrology and mathematics. According to references found in the letters of Synesius, one of her students, Hypatia improved the astrolabe and invented a hydrometer. She also edited and commented on many works of mathematics, including Apollonius' *Cones* and Diophantus' *Arithmetica*, making them more accessible and thus contributing to their survival through the centuries. In 415 Hypatia was murdered during the clashes between followers of Bishop Cyril and the civil governor Orestes, who was an old student of hers.

In this fragment of The School of Athens, *by Rafael, Pythagoras appears in the foreground, and in the background, in a white robe is Hypatia of Alexandria.*

edition contained the Greek text beside the Latin translation, and also contained a whole series of notes and comments.

Bachet's translation gave an extraordinary impetus to the theory of numbers. Bachet himself resolved the linear Diophantine equations $ax + by = cz$. Later Albert Giraud perfectly characterised integer numbers as the sum of two squares. Finally, Fermat invented a new general demonstration method called infinite descent, which he successfully used to demonstrate his famous theory in the case of $n = 4$.

Before Bachet's translation, the theory of numbers was considered to be of little interest to mathematicians. It was thought that the problems they resolved were simply mathematical curiosities, eloquent but rather peculiar and impractical. The main specialities at that time were geometry and analysis. But following Fermat's work the status of number theory made an about turn and became the principal interest for the most outstanding mathematical geniuses: Viète, Fermat, Descartes, Gauss, Euler, Jacobi, Lagrange, Legendre, Dirichlet, Dedekind, Kronecker, etc. – just a few of the mathematicians who researched the innermost characteristics of numbers, which Gauss was to describe as the queen of mathematics.



A portrait of 18th century mathematician Joseph Louis de Lagrange, who studied several problems set by Fermat.

SOLUTIONS TO THE LINEAR DIOPHANTINE EQUATIONS

Diophantine equations require integer solutions to be found for equations with integer coefficients. Linear Diophantine equations were the first to be solved, and they allow solutions to be found to many practical problems. Let's have a look at an example. Let's suppose our housemate goes shopping in order to get a year's supply of oil. He tells us that he has found two offers for bottles of oil, one at £3.24 per litre and the other at £4.50 and that in total he has spent £41.54. We take a look at the box of bottles and tell him that 11 litres of oil will not be enough for the whole year. How were we able to know how many bottles were in the box, if the box was closed? x is assigned to the number of bottles at £3.24 and y to the number of bottles at £4.50; for an equation for the total of pounds spent we get $3.24 \cdot x + 4.50 \cdot y = 43.20$. This equation contains decimals, but multiplying both sides by 100 gives an equation with integer cooefficients: $324 \cdot x + 450 \cdot y = 4{,}320$. Therefore, the problem is solved by determining the values of x, y which satisfy the equation. In fact, we are looking for two integer values, as we know we have bought an integer number of bottles of each type.

The condition that is necessary and sufficient so that a linear equation with integer coefficients has integer solutions is that the maximum common divisor of the coefficients of the unknowns exactly divides the independent term. The maximum common divisor of 324 and 450 is 18, which is an exact divisor of 4,154. Dividing both sides of the equation by 18, gives $18 \cdot x + 25 \cdot y = 240$. A table of values can now be drawn up for resolving the equation. In order to do this, x is given an integer value starting with 0, and the value corresponding to y which satisfies the equation is found, in other words $y = (240 - 18 \cdot x) / 25$.

x	y	x	y
0	9.6	7	4.56
1	8.88	8	3.84
2	8.16	9	3.12
3	7.44	10	2.4
4	6.72	11	1.68
5	6	12	0.96
6	5.28	13	0.24

In this table it can be seen that the only positive integer solutions are $x = 5$, $y = 6$, and that, therefore, our friend has bought 11 litres of oil.

With time, methods of solving this type of equation have been perfected and are now implemented in computer programs and scientific calculators.

In 1885, Sir Thomas Heath published the first English translation of *Arithmetica.* It was an excellent edition that was revised in 1910 and included comments made by Bachet, Fermat and others. To comment on a work was common practice among authors in those times, and the different editions or translations frequently included comments from the editor or translator, although it was not always explicit what belonged to the original and what was a personal contribution from the person who was writing. Perhaps it was taken as given that a work was something that was built on through time and that everyone was invited to understand it, study it and complete it with new contributions. This is why it is so important from a historical point of view to have as many editions as possible, to see how the work has evolved over time.

From the study of the manuscripts currently in existence, Tannery suggested that all of them come from a common source. It seems that this common source could be the edition of *Arithmetica* commented on by Hypatia of Alexandria. According to this theory, the six books that have survived would belong to this edition; any lost books would be from versions that Hypatia did not comment on. If that is true, the passing on of this part of Diophantus' legacy is owed to Hypatia; it is also very probable that this part includes her contributions. This theory is still being researched, and a conclusion has still not been reached.

The problems from Diophantus' *Arithmetica*

One of the copies of the editions commented on by Bachet fell into Fermat's hands. Fermat knew Latin and Greek perfectly, so he could read the work in both languages. Moreover, this was an annotated edition. It was a perfect starting point for him to add his own comments.

Problem 32 of Book II

This problem contains the following statement:

> "Find three numbers such that the square of each of them added to the next gives a square."

To solve this a whole series of tests must be done and with a little luck maybe a solution will be found. We could start, for example, by choosing 1 as the first number. Now, as the statement says, it is squared and the following number is added

to it to give another square. For example, $1^2 + 3 = 4 = 2^2$. 1 and 3 have now been chosen. Now 3 is squared and added to another number to give another square. For example, $3^2 + 7 = 16 = 4^2$. Now we have 1, 3 and 7. Now it is just a case of completing the cycle and checking if squaring the 7 and adding 1 also gives a square: $7^2 + 1 = 50$. What a shame, 50 is not a square number! Therefore, we have to start again, trying with other numbers. The problem is like a jigsaw puzzle; all the pieces have to fit together. Fermat would spend hours feasting on this kind of problem; for him it was a challenge to the imagination, the same type of challenge that he later proposed to his contemporaries.

The solution to problem 32

Diophantus had a solution for the previous problem, and it does not look like he came across it by chance! It was more like he had a magic method of obtaining it. The solution proposed by Diophantus was the following:

> "Let the first be x, the second $2x + 1$ and the third $2 \cdot (2x + 1) + 1$ or $4x + 3$, so that two of the conditions are met. The last condition gives $(4x + 3)^2 + x$ = square number $= (4x - 4)^2$. So $x = 7/57$, and the numbers are 7/57, 71/57, 199/57."

But how is it possible that all this works? Diophantus' talent was undoubtedly extraordinary. He called the first number x. He could have called the second number many things, but he decided to call it $2 \cdot x + 1$ because he knew that $x^2 + 2 \cdot x + 1 = (x + 1)^2$, and therefore the first condition was already met. He could have then called the third number whatever he wanted, but he decided to call it $2 \cdot (2 \cdot x + 1) + 1$, or, $4 \cdot x + 3$, because he knew that $(2 \cdot x + 1)^2 + 2 \cdot (2 \cdot x + 1) + 1 = (2 \cdot x + 2)^2$, and therefore the second condition was already met. Now only the third condition had to be met, namely: $(4 \cdot x + 3)^2 + x$ = square number. And this is where a new touch of genius appears, as Diophantus thought that this square could have the form $(4 \cdot x - 4)$, because then the problem could easily be solved by resolving a very simple equation.

$$(4 \cdot x + 3)^2 + x = (4x - 4)^2.$$

Expanding gives:

$$16 \cdot x^2 + 24 \cdot x + 9 + x = 16 \cdot x^2 - 32 \cdot x + 16.$$

Subtracting $16 \cdot x^2$ from the two sides of the equation gives:

$$24 \cdot x + 9 + x = -32 \cdot x + 16.$$

Resolving the x finally gives that

$$\begin{gathered} 24 \cdot x + x + 32 \cdot x = 16 - 9 \rightarrow \\ 57 \cdot x = 7 \rightarrow \\ x = 7/57, \end{gathered}$$

which is the first value we were looking for. From there it is easy to obtain the second number, $2 \cdot x + 1 = 71/57$, and the third number, $4 \cdot x + 3 = 199/57$. Finally, it can easily be proved that:

$$\begin{array}{ll} (7/57)^2 + 71/57 = 4{,}096/3{,}249 = (64/57)^2 & \text{(first condition);} \\ (71/57)^2 + 199/57 = 16{,}384/3{,}249 = (128/57)^2 & \text{(second condition);} \\ (199/57)^2 + 7/57 = 40{,}000/3{,}249 = (200/57)^2 & \text{(third condition).} \end{array}$$

Characteristics of the problem

In this problem Diophantus' style can be appreciated, and Fermat must have marvelled at it. On the one hand there was the type of problem he resolved: aesthetically beautiful, but apparently useless. Who would be interested in solving this problem? It is no use for counting crops, or measuring land, or positioning stars... it is simply a demonstration of the properties of rational numbers. Its use is found in the symphony of numbers, in the eternal quest for understanding the internal harmony and amalgamated rhythms. However, resolving it is such a profound challenge for the mind that it tests all mathematical machinery and its instruments. In fact, it was Viète who, pondering *Arithmetica*, established the basis of the algebraic notation which is used today; his intention was to make Diophantus' work more understandable and to resolve more and more complex problems. Also Fermat, starting with *Arithmetica*, came up with new problems and

invented new methods of demonstration that reignited interest in the theory of numbers. For their part, prime numbers, which so concerned the Greeks, are the base of the most sophisticated systems that are used today to encrypt information and for modelling the universe.

On the other hand, the resolved problem is purely arithmetic. If it had a geometric aspect, adding a number to the square with another number to give a resulting square, it would be, for example, adding an area to a length, to give an area, which would be adding geometric magnitudes from different dimensions. The theorem of Pythagoras is different, because it states that the square constructed on the sides total the square constructed on the hypotenuse; in other words, two areas added together give another, larger area. Geometrically, the magnitudes involved in the equation are the same. Also, in Fermat's theorem the magnitudes involved are of the same degree, $x^n + y^n = z^n$. For the exponent $n = 3$ it could be imagined that two volumes are being added together to give one larger one. For larger exponents we would be talking about hyper-volumes in higher-dimensional spaces.

Parallel reasoning

Another characteristic of the problem is that the solution is far from trivial. It does not appear to be a mere coincidence. This characteristic is found in many of the problems in *Arithmetica*. It also seems that Diophantus is happy to find just one particular solution without worrying about resolving the general problem and so finding all the possible solutions. However, his digressions open the doors to a universe of parallel reasoning that allows new solutions to be found that do not appear in the book.

For example, if instead of making the last condition equal $4x - 4$ it had been made to equal $4x - 5$, another, perfectly valid solution would have been obtained:

$$
\begin{gathered}
(4 \cdot x + 3)^2 + x = (4 \cdot x - 5)^2 \rightarrow \\
16 \cdot x^2 + 24 \cdot x + 9 + x = 16 \cdot x^2 - 40 \cdot x + 25 \rightarrow \\
24 \cdot x + 9 + x = -40 \cdot x + 25 \rightarrow \\
24 \cdot x + x + 40 \cdot x = 25 - 9 \rightarrow \\
65 \cdot x = 16 \rightarrow \\
x = 16/65.
\end{gathered}
$$

Therefore, we get another solution, consisting of the numbers 16/65, 97/65, 259/65.

If, instead of making the last condition equal $4x - 4$ it had equalled $5x - 3$, it would also have given another, perfectly valid solution:

$$(4 \cdot x + 3)^2 + x = (5 \cdot x - 3)^2 \rightarrow$$
$$16 \cdot x^2 + 24 \cdot x + 9 + x = 25 \cdot x^2 - 30 \cdot x + 9.$$

Now, subtracting the 9 from both sides of the equation gives:

$$16 \cdot x^2 + 24 \cdot x + x = 25 \cdot x^2 - 30 \cdot x.$$

Dividing both sides by x gives:

$$16 \cdot x + 24 + 1 = 25 \cdot x - 30 \rightarrow$$
$$24 + 1 + 30 = 25 \cdot x - 16 \cdot x \rightarrow$$
$$55 = 9 \cdot x \rightarrow$$
$$x = 55/9.$$

Therefore, we have another solution with the numbers 55/9, 119/9, 247/9. Also, they open the door to new challenges. For example, are there integer solutions that also meet these conditions?

Problem 29 of Book IV

One of the problems from *Arithmetica* with the most history is problem 29 of book IV. It states:

> "Find four square numbers such that their sum added to the sum of their sides makes a given number."

Again we can see Diophantus' genius:

> "Given the number 12. Now $x^2 + x + 1/4$ = a square number. Therefore the sum of four squares + the sum of their sides + 1 = the sum of four other squares = 13, by hypothesis. Therefore we have to divide 13 into [the sum of] four squares; then, if we subtract 1/2 from each of their sides, we shall have the sides of the required squares. Now, 13 = 4 + 9 = (64/25 +

36/25) + (144/25 + 81/25), and the sides of the required squares are 11/10, 7/10, 19/10, 13/10, the squares themselves being 121/100, 49/100, 361/100, 169/100."

Impeccable reasoning for the particular case of $n = 12$. The problem set by Diophantus, in modern notation, could be expressed as follows:

"Find x_1, x_2, x_3, x_4 so that

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_1 + x_2 + x_3 + + x_4 = n,$$

where n is a given number".

Adding 1 to each side of the equation gives:

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_1 + x_2 + x_3 + x_4 + 1/4 + 1/4 + 1/4 + 1/4 = n + 1.$$

Reordering the terms and assuming that $n = 12$ gives:

$$x_1^2 + x_1 + 1/4 + x_2^2 + x_2 + 1/4 + x_3^2 + x_3 + 1/4 + x_4^2 + x_4 + 1/4 = 12 + 1.$$

Taking into account that $x^2 + x + 1/4 = (x + 1/2)^2$, it can be written:

$$(x_1 + 1/2)^2 + (x_2 + 1/2)^2 + (x_3 + 1/2)^2 + (x_4 + 1/2)^2 = 13.$$

Therefore, we need to express 13 as the sum of four square numbers. In this particular case it is easy to see that 13 is the sum of two square numbers, 4 + 9, and the Pythagoras theorem can be used to express each of these squares as a sum of two others, as Diophantus himself explained in another of the problems in *Arithmetica*.

The numbers 4, 3, 5 are a Pythagorean triple, in other words $4^2 + 3^2 = 5^2$; from there, dividing by 5^2, we get $(4/5)^2 + (3/5)^2 = 1$. Now, if we multiply both sides of the equation by 2^2, we get $(8/5)^2 + (6/5)^2 = 2^2$, in other words, $(64/25) + (36/25) = 4$; and if we multiply both sides of the equation by 3^2, we get $(12/5)^2 + (9/5)^2 = 3^2$, in other words, $(144/25) + (81/25) = 9$, which is the decomposition proposed by Diophantus. Therefore, we get the solution:

$$(x_1 + 1/2) = 8/5,$$
$$(x_2 + 1/2) = 8/5,$$
$$(x_3 + 1/2) = 8/5,$$
$$(x_4 + 1/2) = 8/5.$$

Subtracting 1/2 we get the solution proposed by Diophantus. The incredible thing about this solution is that $13 = 1 + 4 + 4 + 4$, in other words, there was no need to go to such lengths to break down 13 as the sum of four squares! This breakdown would give the solution 1/2, 3/2, 3/2, 3/2.

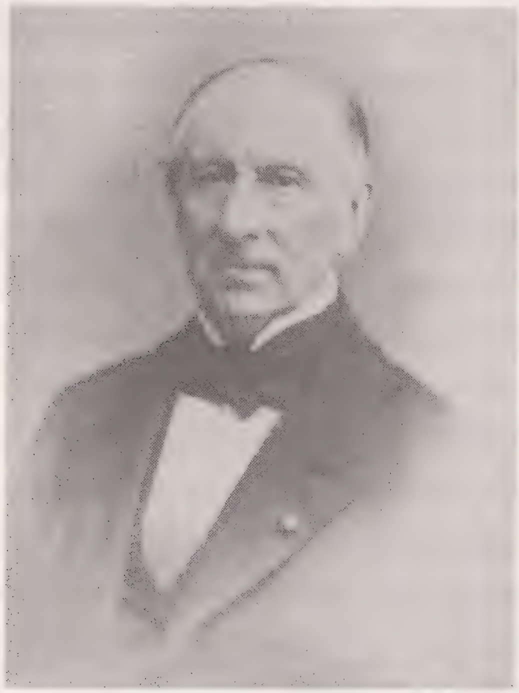
An enigmatic annotation

Bachet realised that Diophantus seemed to assume in this, and other problems from *Arithmetica*, that all numbers can be expressed as a sum of four squares, adding that he himself had verified this fact for numbers up to 325, but he would like to have a proof. Then the genius of Fermat came into play and he wrote: "I have been the first to discover a most beautiful theorem of the greatest generality, namely this: every number is either a triangular number or the sum of two or three triangular numbers; every number is a square or the sum of two, three, or four squares; every number is a pentagonal number or the sum of two, three, four, or five pentagonal numbers; and so on ad infinitum, for hexagons, heptagons, and any polygons whatsoever, the enunciation of this general and wonderful theorem being varied according to the number of angles." And he finished by saying: "The demonstration of this, which depends on several intricate mysteries of numbers, I cannot provide here. I have decided to dedicate a separate and complete work to this subject, thus extraordinarily advancing arithmetic in this area of research beyond previously recognised limits."

But this work never saw the light of day. Did he ever write it? Did he really have a demonstration? It is not known. This is another of Fermat's enigmas. What is certain is that mathematicians such as Legendre, Lagrange, Euler and Gauss dedicated themselves to thinking about this problem and provided partial solutions to the issue.

In 1770 Joseph Louis de Lagrange demonstrated the case of the square, in other words, that all natural numbers can be expressed as a sum of four square numbers. The demonstration of the triangular case is owed to Gauss, who on 10 July wrote in his diary: **EΥRHKA num$=\Delta+\Delta+\Delta$.

This case was the equivalent of demonstrating that all numbers of the form $8 \cdot m + 3$ can be expressed as the sum of three odd squares. Dirichlet, for his part, dedicated himself to the study of how many ways a given number can be expressed as the sum of three triangular numbers. Finally, in 1813, Cauchy found a complete demonstration. It took nearly 150 years to resolve a note in the margin.



A portrait of mathematician Augustin Louis Cauchy, who completed the demonstrations suggested by Fermat of problem 29 of Diophantus' Book IV.

Back to Book II: Problem 8

Problem 8 of Book II is undoubtedly essential in this story. In it Diophantus sets the following problem:

"Divide a square into the sum of two squares."

He then gives the following solution:

"Let the square number be 16. Let x^2 be one of the square numbers we are looking for. Thus $16 - x^2$ will be equal to a square number. Take a square number of the form $(m \cdot x - 4)^2$, m being any integer and 4 the number

which is the square root of 16. Take $(2 \cdot x - 4)^2$ as an example, and make it equal $16 - x^2$. Thus $4 \cdot x^2 - 16 \cdot x + 16 = 16 - x^2$, or $5 \cdot x^2 = 16 \cdot x$, and $x = 16/5$. Thus the squares we are looking for are 256/25, 144/25."

The technique is the same as that used in problem 32 of Book II. Any other number could have been used as m, giving rise to an infinity of solutions, all of them easy to find. In the margin of this page dedicated to this problem Fermat wrote a comment which would make history:

Cubum autem in duos cubos, aut quadrato-quadratum in duos quadrato-quadratos, et generaliter nullam in infinitum ultra quadratum potestatem in duos eiusdem nominis fas est dividere cuius rei demonstrationem mirabilem sane detexi. Hanc marginis exiguitas non caperet."

Its translation is as follows:

"It is impossible for a cube to be the sum of two cubes, a fourth power to be the sum of two fourth powers, or in general for any number that is a power greater than the second to be the sum of two like powers. I have discovered a truly marvellous demonstration of this proposition that this margin is too narrow to contain."

In other words, it states that the equation $x^n + y^n = z^n$ has no rational solutions when $n > 2$, and Fermat gives the excuse that the marvellous demonstration that he had in his head did not fit in the margin of the page. This excuse about a lack of space is similar to the note on problem 29 in Book IV. Of course, his long-awaited solution never saw the light of day.

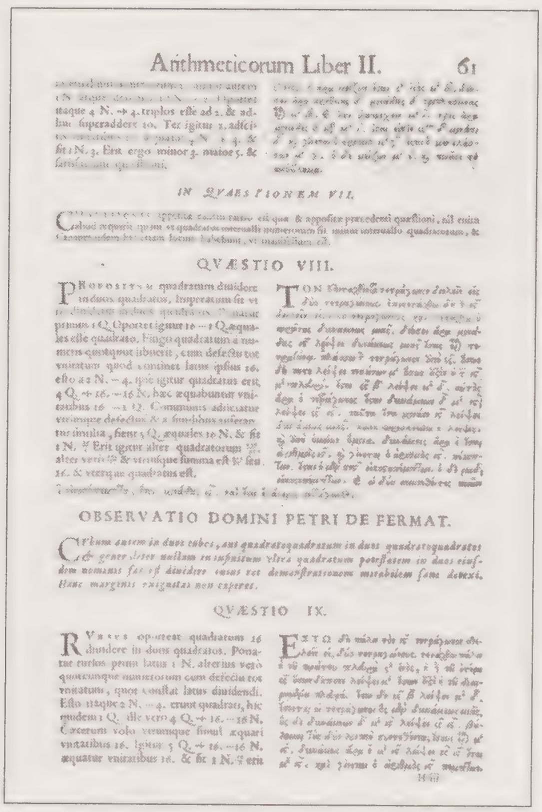
This comment, and others of Fermat, never cease to amaze. On the one hand, it seemed like he never had the intention of publishing them. That is why no demonstrations of any kind were ever expected from him. They seem more like personal notes that helped him to remember what he had thought and go into the problem in more detail. On the other hand they are written in a style that assumes someone will read them. Otherwise, what sense does it make to explain to himself that he has a marvellous demonstration that the margin of the page does not allow him to provide, or that he cannot provide a demonstration because he has decided to publish a complete work on the matter? It is as if he used the

book like a personal notepad, but at the same time he had it in mind to prepare a commented edition of *Arithmetica* with the intention of publishing it one day.

Fermat's contributions

Whatever his intentions, the comment was not lost. Fermat returned to it on several occasions, and actually tried to write down his "marvellous demonstration". The first thing Fermat realised was that all rational solutions give rise to an integer solution by multiplying by the least common multiple of the denominators. Therefore, it is sufficient to demonstrate that the equation does not have integer solutions.

It is also easy to see that only the cases for $n = p$, where p is a prime number, and $n = 4$ need be demonstrated. All other cases would be demonstrated automatically. Given that if $n = p \cdot m$, then the equation $x^n + y^n = z^n$ becomes $x^{m \cdot p} + y^{m \cdot p} = z^{m \cdot p}$, of which leads to $(x^m)^p + (y^m)^p = (z^m)^p$. If there is no solution for exponent p, there will be none for a multiple of p. In the same way it is evident that if there is no solution for exponent $n = 4$, then nor will there be for exponents that are multiples of 4. So Fermat focused on demonstrating that his equation had no integer solutions


Arithmeticorum Liber II. 61

IN QVAESTIONEM VII.

QVÆSTIO VIII.

OBSERVATIO DOMINI PETRI DE FERMAT.

Cubum autem in duos cubos, aut quadratoquadratum in duos quadratoquadratos & generaliter nullam in infinitum vltra quadratum potestatem in duos eiusdem nominis fas est diuidere cuius rei demonstrationem mirabilem sane detexi. Hanc marginis exiguitas non caperet.

QVÆSTIO IX.

The page of Book II of Diophantus' Arithmetica *with problem 8 in the 1670 edition, which includes Fermat's comment.*

for $n = p$, where p is a prime number and the case where $n = 4$.

Everything seems to indicate that he had resolved the cases $n = 3$ and $n = 4$. The demonstration of the case $n = 3$ has never been found but Fermat makes reference to it in some of his letters. That of $n = 4$ has been found, and it is magnificent. In it he invents the method of infinite descent. He shows that there are three natural values x, y, z other than zero that verify the equation $x^4 + y^4 = z^4$, so three other, smaller natural values other than zero x', y', z' can also verify the same equation. In this way, with successive reasoning, smaller and smaller solutions, all of them natural and not zero, would be obtained. But this clearly leads to an absurdity because natural numbers cannot be infinitely small. Therefore, he concludes that these solutions do not exist.

With his demonstration of infinite descent, Fermat believed he had devised a general method that allowed any question related to the theory of numbers to be resolved, as Descartes thought that with the help of analytical geometry he could resolve all problems in nature. But reality always turns out to be more evasive. Its

FERMAT'S PARTIAL SOLUTIONS FOR THE 3RD DEGREE

Although it is true that the equation $x^3 + y^3 = z^3$ has no integer solutions other than 0, the truth is that it "almost" does, because there are some values of x, y, z that almost fit the equation. It is quite easy to see that $5^3 + 6^3 = 7^3 - 2$ only differs by two units from the equation suggested by Fermat. A more striking case would be that of $6^3 + 8^3 = 9^3 - 1$. It really seems incredible that, being so close to a solution, there are no integers that fit!

What would happen if a term was added to Fermat's equation? It is surprising to find that in that case it does have solutions other than zero! In effect, $3^3 + 4^3 + 5^3 = 6^3$, and $7^3 + 14^3 + 17^3 = 20^3$.

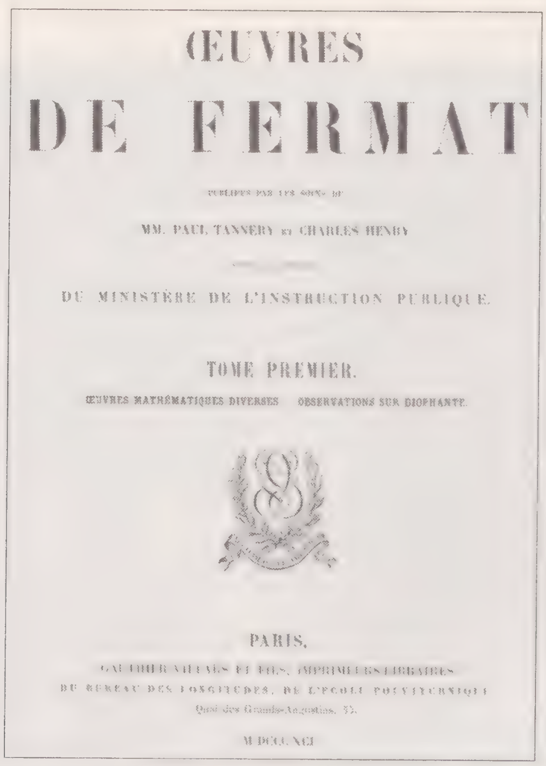
With its characteristic sense of humour, one of the scenes from *The Simpsons* shows the equation $1{,}782^{12} + 1{,}841^{12} = 1{,}922^{12}$. Could Lisa have finally solved Fermat's enigma? With further analysis it can be seen that this is a case of an "almost" solution, which coincides exactly until the ninth digit of the number written in base ten. In another episode she further improves her solution; thus, in *The Wizard of Evergreen Terrace* $3{,}987^{12} + 4{,}365^{12} = 4{,}472^{12}$, an "almost" solution that fits until the tenth figure and which also coincides in the number of the units of the number written in base ten. A standard calculator with an 8-digit screen would not even detect this result.

variety does not allow it to be tamed by one unique method, however powerful. New exceptions always come up that once again test human genius, and they demand that we always better ourselves in order to reach higher – and more profound – levels. And that is exactly what happened with Fermat's last theorem.

Infinite descent undoubtedly allowed Fermat to obtain a demonstration for the case of $n = 3$, but he probably realised that superior cases were not going to be tamed in the same way. Even so, Fermat's contribution was spectacular. Because, in demonstrating the case $n = 4$ he had uncovered a new method for mathematical demonstrations that turned out to be extraordinarily fruitful. And he had also demonstrated his theory for half of the exponents, of which there are many! But at the same time he left the resolution of all other cases open, a task to which the most eminent mathematicians of every era would unsuccessfully dedicate themselves.

An unpublished genius

It has been said many times that Fermat did not want to publish his work. But it is not necessarily true. In fact, as early as 1636 he sent Mersenne his *Maximums and Minimums Method*, and asked him to make it known to the mathematicians of Paris.


ŒUVRES

DE FERMAT

MM. PAUL TANNERY ET CHARLES HENRY

DU MINISTÈRE DE L'INSTRUCTION PUBLIQUE.

TOME PREMIER.

ŒUVRES MATHÉMATIQUES DIVERSES — OBSERVATIONS SUR DIOPHANTE.

PARIS,

Quai des Grands-Augustins, 55.

MDCCCXCI

An 19th-century illustrated edition of the work's of Fermat's as published after his death.

Also, in his correspondence, throughout his life, he not only set challenges, but also gave clues to the solutions, and in some cases explained some of his methods in great detail.

In 1654, Fermat re-initiated his correspondence with mathematicians in Paris. Blaise Pascal got in contact with him to ask his opinion on his ideas about probability, and Fermat responded brilliantly regarding combinatorial analysis. Through his letters he was establishing the basis of a new discipline of mathematics: the theory of probability, and Fermat made the most of the occasion by making some of his most recent results known.

Firstly, Fermat set new challenges on the theory of numbers to Blaise Pascal, Gilles Personne de Roberval, John Wallis, William Brouncker, Bernard Frénicle de Bessy, etc. Among the challenges he proposed were finding the integer solutions to $Nx^2 + 1 = y^2$, where N is not a square number; demonstrating that the equation $x^2 + 2 = y^3$ only has one solution with natural numbers; and demonstrating that the equation $x^2 + 4 = y^3$ only has two natural solutions. Fermat was aware that in order to advance these and other similar problems he needed to hone new methods of demonstration that were to advance mathematics profoundly. However, his contemporaries saw these challenges as very peculiar cases that were not worth dedicating much time to. However, little by little, Fermat managed to

William Brouncker was one of the many mathematicians with whom Fermat maintained a correspondence.

convince some of his correspondents to dedicate some of their time to studying these problems.

Fermat then asked Pascal and De Carcavy to help him find an editor for his work *Novus Secundarum et Ulterioris Radicum in Analyticis Usus*. He wrote to De Carcavy on 9 August 1654, regarding this matter: "If it is not to your dislike, the two of you [De Carcavy and Pascal] may undertake that publication [of the work], of which I consent to your being the masters, you may clarify or supplement whatever seems too concise and relieve me of a burden that my duties prevent me from taking on. I would like the work to appear without my name, and I leave it in your hands to attribute authorship to whoever you consider a friend." In turn De Carcavy contacted Huygens. But the fact is that none of them managed to get it to publication. So time went by and Fermat's work was threatened to be forgotten forever.

Fermat, aware of this situation, accommodated it in the best way he knew – by writing more letters. He did not want his name to appear in his work; the important thing for him was the legacy he left to science so that it could continue to advance. In 1659 he asked De Carcavy to send Huygens his *Relation des Nouvelles Découvertes en la Science des Nombres*. This work explained many of his methods in more detail than was customary but not with all the detail his colleagues would have liked.

THE $4k + 1$ CASE

An example of the not entirely complete explanations that appear in *Relation des Nouvelles Découvertes en la Science des Nombres* are the clues provided by Fermat to demonstrate, using the method of infinite descent, that all prime numbers of the type $4k + 1$ can be expressed as the sum of two squares. Starting with the assumption that there is a prime number of the type $4k + 1$ which cannot be expressed as the sum of two squares, it is then affirmed that there would be a smaller prime number of the same type which could not be expressed as the sum of two squares either; which leads to a contradiction, given that the equation eventually results in the number 5, the smallest of the prime numbers of this type, and which it turns out can be expressed as the sum of two squares: $5 = 2^2 + 1^2$. Therefore, the proposition is proven. But Fermat does not explain how to obtain the smallest prime number by starting with the largest. It would be several decades until Euler completed the step which Fermat left unexplained. Another of the margins left empty when Fermat had the answer in his mind, perhaps?

Fermat died before he could find an editor to publish his work. Although some of it came to light through his letters or the publications of other contemporary mathematicians who knew of his results, the truth is that most of them probably would have been lost if it were not for the efforts of his son Samuel, with whom he shared a passion for mathematics. In 1670, Samuel Fermat managed to publish an edition of Bachet's translation of Diophantus' *Arithmetica* with his father's observations.

The methods proposed by Fermat surprised his contemporaries over and over again. History has demonstrated that his contributions were at the origin of an almost endless list of new theories that were established during that period. To resolve problems of maximums and minimums he invented an equivalent to the derivative equalled to zero. He solved problems using geometry before Descartes, and could be considered the father of analytical geometry. He resolved probability and combinatorial problems, and is considered, together with Pascal, one of the founders of these new theories. Fermat is also considered the father of the modern theory of numbers, where the heights of his excellence and the value of his contributions were sublime. But there is also the Fermat who investigated optics and mechanics. As with Midas who turned everything to gold, any problem Fermat touched resulted an essential contribution to the advancement of science. All this was done in his spare time between legal cases. If he had thought of trying to solve the equation for justice, maybe he could have found the solution.

Chapter 5

The Ingredients for a Tasty Dish

There is no problem that can withstand the assault of sustained thought
Voltaire

In 1666, a few years after the death of its great promoter, Mersenne, the Académie des Sciences was founded in Paris. Jean Baptiste Colbert, minister of finance at the time, had granted ample resources for this prestigious institution. Little by little he invited eminent scientists from the world over to be a part of it, among them were many of Mersenne's correspondents. In fact, it was this select group of scientists that took responsibility for driving the ambitious project and making it a reality.

Fermat's Grand Prix

In 1721 the Académie des Sciences de Paris established a series of prizes in order to stimulate scientific advances on specific subjects of significant importance. The board of judges was formed by experts of world recognised prestige. Among the winners were researchers such as Colin Maclaurin for his work on the impact of bodies in 1724, Pierre Bouguer and Charles Étienne Louis Camus for their contributions to the design of boat masts in 1727, Leonhard Euler for a study on the nature of fire in 1738, Charles Augustin de Coulomb for his research on friction in 1781, Siméon Denis Poisson for his results on electricity and Augustin Jean Fresnel for his studies on refraction in 1812.

At that time the challenges that Fermat had set for the world were a source of despair for numerous mathematicians, who tried to prove each of the statements, sometimes successfully and sometimes not. The most stubbornly resistant conjecture of all began to be known as 'Fermat's last theorem'. In 1826, the Académie, with the intention of making a breakthrough and stimulating research, dedicated an edition

of the prize to anyone who could demonstrate the last theorem. Many scientists dedicated themselves to this work, or encouraged their colleagues to do so.

The first two hundred years

Heinrich Wilhelm Matthäus Olbers was a doctor and astronomer, who dedicated long hours to observing the sky. In 1802 he had located the giant asteroid Ceres in the place where Gauss had predicted it would be, one year after Giuseppe Piazzi discovered the object before later losing sight of it. In 1807 Olbers discovered a second asteroid and gave Gauss the honour of naming it. The name proposed by Gauss was Vesta, the Roman god of the home. Vesta turned out to be the brightest body in the asteroid belt, oc-

Monarch Louis XIV's visit to the Académie des Sciences in Paris in 1671, depicted in an engraving by Sébastien Le Clerc included to the volume *Mémoires pour Servir a l'Histoire Naturelle des Animaux.*

casionally visible by the naked eye from the Earth, a sixth magnitude star. A few thousand years ago Vesta lost 1% of its mass as a result of an impact, and many of the fragments from it hit Earth in the form of meteorites. One of the problems considered by Olbers was why the night sky was so dark if it was lit by an infinity of stars. Surely they should make it a bright as daytime. Now this idea is called 'Olber's Paradox'.


German doctor and astronomer Heinrich Olbers according to a lithograph by Rudolf Suhrlandt.

When he found out about the Académie's prize for a proof of the last theorem, Olbers contacted his friend Carl Friedrich Gauss and urged him to enter. On 21 March 1816, Gauss responded saying: "I must confess that Fermat's theorem as an isolated proposition is of little interest to me, because I could easily establish a multitude of propositions of this type that could not be demonstrated or refuted." However, Gauss also dedicated his time to thinking about the theorem, as demonstrated by his personal notes, which contained proofs for $n = 3$ and $n = 5$. Gauss may have already made these attempts when Olbers proposed it to him and, aware of the difficulty, preferred to turn down the invitation and continue investigating in private in the hope of finding a result. Whether he actually dedicated only a little time to this problem because he preferred to focus his efforts on subjects that he considered to be more interesting remains a mystery.

Despite Gauss' words, the truth is that the great mathematicians of the time were spurred on by the theory and fervently dedicated themselves to the search for a proof. It was no longer just the Académie's prize at stake, whoever could demonstrate it would be bestowed with fame and glory. The deadline came for the submission of results, and nobody had managed to come up with a proof! It is probable that the Académie was expecting this. Before the prize was established, the likes of Euler had already taken the problem on, and even a genius of his magnitude only managed to demonstrate it for $n = 3$, in around 1760. As was already mentioned in the previous chapter, it is probable that Fermat already had a demonstration of this exponent using

Portrait of German mathematician Johann Peter Gustav Lejeune Dirichlet.

his method of infinite descent, but the mathematical community could sleep more peacefully knowing that Euler had it written down. It was clear that a cubic number could not be broken down into the sum of two cubic numbers, but could the same be said of the exponents that remained?

Fascination about the last theorem was growing in the scientific community. German Johann Peter Gustav Lejeune Dirichlet (1805-1859) and Frenchman Adrien-Marie Legendre (1752-1833) shared a few moments of glory when they independently managed to provide proof for the exponent $n = 5$, in around 1825. In 1832 Dirichlet managed to take one step more and demonstrated the last theorem for the exponent $n = 14$. In 1839 Frenchman Gabriel Lamé (1795-1870), eight years before announcing his failure to find a general proof, won his place in history by demonstrating the case $n = 7$. Each mathematician managed to demonstrate a few exponents. Taking into account that there are infinite prime exponents, would it take an infinity of lives to find a demonstration?

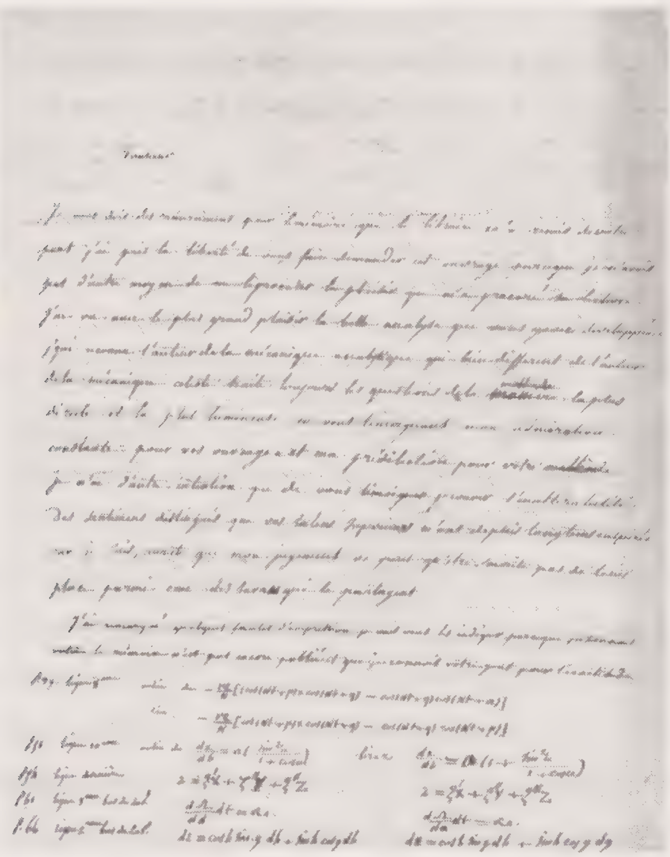
An unexpected protagonist

The first sign of hope that many cases could be covered at the same time came at the hands of Marie-Sophie Germain (1776-1831), perhaps the greatest female mathematician of all time, who in 1823 demonstrated that if p and $2p + 1$ are two primary numbers greater than 2, then $x^p + y^p = z^p$ does not have primitive solutions (in other words, no common factors) in which xyz is not divisible by p. The Académie's regulations did not allow women to submit their results personally, so it was Legendre and his colleague Auguste-Louis Cauchy who took it upon themselves to inform the scientific community.

As we saw in the previous chapter, if it could be proven for the prime exponents, then it would be proven for all natural exponents. Equally, it is easy to see that if an integer solution x, y, z has a common factor, dividing by that factor also gives an

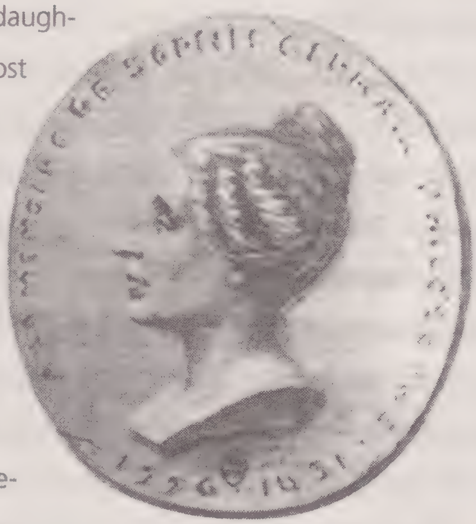
integer solution. Therefore, a demonstration for the theorem for primal solutions is also a general proof of it. After Germain, two cases from all the solutions to the last theorem began to stand out. The first case consisted of the solutions such that neither x, nor y, nor z were divisible by p. The second case was formed by solutions such that either x, or y, or z were divisible by p. "In one stroke of the pen," as Legendre would say, Germain's result proved the first case of Fermat's theorem for an enormous group of exponents. Those that remained in order to reach 100 were completed by Legendre himself, who demonstrated that if p is a prime number such that $4p + 1$, or $8p + 1$, or $10p + 1$, or $14p + 1$, or $16p + 1$, are prime numbers, then the first case of Fermat's theorem was demonstrated for exponent p. It was not until 1977 that Terjanian demonstrated that the first case of the theorem is verified for all even exponents $2p$, where p is a prime number.

If, for example, the exponent $p = 5$ is considered, it can be observed that $2p + 1 = 11$, which is also a prime number; therefore, from Germain's work, the first case of Fermat's theorem is demonstrated for this exponent. On the other hand, p



A letter from Marie-Sophie Germain to mathematician Joseph Louis de Lagrange. The French mathematician's research represented an important step forward in resolving Fermat's last theorem.

SOPHIE'S CHOICE



Marie-Sophie Germain.

Maria-Sophie Germain was born in Paris in 1776, the daughter of a prosperous silk merchant whose house was host to philosophical and political discussions. At the age of three, Sophie read the well-known anecdote about the death of Archimedes at the hands of a Roman soldier and, she was so moved by it that she herself decided to become a mathematician. At the outbreak of the French Revolution, her parents kept her hidden in the house for seven or eight years in order to protect her. The young girl made the most of that period and her father's library to study mathematics on her own. Sophie read Newton and Euler in secret at every opportunity. Sophie's decision to dedicate her life to an academic career, something practically unheard of at the time, was unshakeable, and her family had no choice but to accept it. At 18 years of age, and faced with the bar on women signing up to the recently established École Polytechnique in Paris, where leading figures such as Lagrange taught, Germain posed as an older student friend of the family, Antoine Auguste Le Blanc, in order to be able to attend. Under that pseudonym she submitted some work to Lagrange, who was so impressed that he requested an interview. Sophie had no choice but to reveal her true identity, and Lagrange, despite his initial surprise, offered to be her professor, which introduced the young woman to many Parisian scientific circles.

Germain maintained a fruitful correspondence with Gauss, always under her pseudonym. When he found out the true identity of his correspondent in 1806, the mathematician wrote to her: "The taste for abstract science in general, and especially for the mysteries of numbers, is very rare. But a woman, because of her sex and our customs and prejudices, must encounter infinitely more obstacles, and when a person of that sex is nonetheless able to break through these barriers and penetrate the most hidden secrets, she must undoubtedly have the most noble courage, quite extraordinary talent, and superior genius." In 1811 Germain was the only participant in a competition set by the Académie des Sciences to explain the mathematical basis of vibrations in elastic surfaces. Having been rejected on two occasions, in 1816 she won the prize, which made her the first woman to attend the Académie's sessions (other than the members' wives). In 1830 the University of Göttingen agreed to award her an honorary degree, but Germain died the following year, before receiving the title.

$= 7$ gives $2p + 1 = 15$, which is not a prime number. Thus, taking into account Germain's results, this exponent would remain unproven. But as $4p + 1 = 29$ is a prime number, the first case of Fermat's theorem would be demonstrated with Legendre's results.

Lamé's demonstration

On 1 March, 1847, Gabriel Lamé made a spectacular announcement to the Académie des Sciences in Paris. He had found the long-awaited demonstration of Fermat's theorem for all exponents! The French researcher submitted the work that had led him to the coveted solution to the mathematical community. His reasoning was simple, and was based on results found by previous mathematicians. The idea was to work on the body of complex numbers, where the square root of minus one, $\sqrt{-1}$, makes sense and is designated by the letter i. In this group, $x^2 + y^2$ becomes the product of two complex numbers $(x + yi) \cdot (x - yi)$, converting an addition problem to a multiplication one. With this work tool, the right-angle triangle theorem changes from being written:

$$x^2 + y^2 = z^2,$$

to being written:

$$(x + yi) \cdot (x - yi) = z^2.$$

This last equation can now be resolved in the group of complex numbers as follows: $x + yi$, where x, y are integers; x is called the real part and y, the imaginary part (the group is given the name of Gauss integers). This group is very similar to that of integers: in it, adding, subtracting and multiplying can be done without problems. We can also talk of divisibility and of prime numbers. The fundamental theorem of arithmetic is verified in the same way, which states that all numbers can be broken down in one unique way into a product of prime factors (in the case of complex numbers, there are some exceptions). A very interesting consequence of this theorem is that if the product of two co-prime numbers gives a square number, then each of these numbers must in turn be a square number. Following this reasoning, finding a Pythagorean triple would mean finding a primal solution x, y, z to the equation $x^2 + y^2 = z^2$, in other words, a solution where x, y, z have no common factors. A solution of this type means that Gauss numbers $x + yi, x - yi$ do not have Gauss

factors in common either. Thus, two Gauss integer coprimes need to be found such that the product is a square number.

In summary, if there is a primal solution for the equation $x^2 + y^2 = z^2$, then we get a product of two Gauss integer coprimes which give a square number as a result. Therefore, each of the factors will also be a square number. Thus:

$$x + yi = (a + bi)^2 = a^2 + 2abi + (bi)^2 = a^2 - b^2 + 2abi.$$

Now, equating the real part with the real part and the imaginary part with the imaginary part, we get:

$$x = a^2 - b^2$$
$$y = 2ab.$$

This is the formula that already appears in Euclid's *Elementos* for generating Pythagorean triples. The reasoning proposed by Lamé followed the same line of demonstration. Fermat's equation $x^p + y^p = z^p$ was transformed, with the help of complex numbers, to a product of factors. This time the factors would have to contain roots of the degree p of the unit. In the group of complex numbers, in the same way that 1 has two square roots, +1 and −1, p also has roots of degree p which are designated by $1, \zeta, \zeta^2, \zeta^3 \ldots \zeta^{p-1}$. Using these roots, it can be written as:

$$x^p + y^p = (x + y)(x + \zeta y)(x + \zeta^2 y)(x + \zeta^3 y) \ldots (x + \zeta^{p-1} y) = z^p.$$

Therefore the first step, which consisted of translating an additive problem into a multiplication problem, worked.

The following step involved considering Fermat numbers of the form:

$$a_0 + a_1 \zeta + \zeta^2 a_2 + \zeta^3 a_3 + \ldots + \zeta^{p-2} a_{p-2}.$$

These numbers are called cyclotomic numbers. They can be added, they can be subtracted, multiplied and divided and we can also talk of prime cyclotomic numbers. Lamé continued, as in the case of the Gauss integers, until he had demonstrated the theory! Illustrious mathematician Joseph Liouville, who was listening carefully, asked to speak and came up with a question. Was it demonstrated that the factorisation of the group of cyclotomic numbers is unique? Because otherwise, the whole demonstration would fail! Lamé recognised that he had not demonstrated

it, but was sure he could justify it quickly and complete the missing details. Lamé's justification however, never arrived.

Ideal solutions

A few months later, German Ernst Eduard Kummer wrote a letter to Liouville. In it, he explained a flaw in Lamé's attempt at a demonstration, the property of unique factorisation was not verified in general for cyclotomic numbers. For example, it fails in the case of $p = 23$. However, he continued, "it is possible to solve it by introducing a new type of complex numbers, which I have called ideal complex numbers." The ideals introduced by Kummer allowed him to recuperate the unique factorisation and continue to advance the reasoning.

Below are two examples to illustrate Kummer's idea. First, consider the following group of even integers:

$$2\,Z = \{\ \dots -4, -2, 0, 2, 4, 6, 8, 10\ \dots\ \}.$$

This group can be added, subtracted and multiplied without problems. In this group the 10 cannot be broken down as a product of two even numbers, and therefore would be a 'prime number'. Also, the 2 and the 50 would be 'prime numbers'. On the other hand, the 100 can be broken down into two different forms as a product of 'prime numbers':

$$100 = 10 \cdot 10 = 2 \cdot 50.$$

Therefore, in the group of even numbers, the property of unique factorisation is not verified. To rescue this property an 'ideal' number can be introduced, the 5, which does not belong to the group of even numbers. With the help of this number, the 10 and the 50 can now be broken down to leave 'prime numbers', giving:

$$100 = 10 \cdot 10 = 5 \cdot 2 \cdot 5 \cdot 2.$$
$$100 = 2 \cdot 50 = 2 \cdot 2 \cdot 5 \cdot 5.$$

And the two factorisations coincide!

A portrait of German mathematician Ernst Eduard Kummer.

In the second example, proposed by Richard Dedekind in 1870, the group of numbers is considered in the following way:

$$Z\left[\sqrt{5i}\right]=\left\{a+b\sqrt{5i}: a,b\in Z\right\}.$$

In this group, the numbers 2, 3, $(1+\sqrt{5i})$, $(1-\sqrt{5i})$, are prime numbers. The 6, which is not, can be broken down into two different forms as the product of prime numbers:

$$6=2\cdot 3=(1+\sqrt{5i})(1-\sqrt{5i}).$$

Therefore, this group does not verify the property of unique factorisation. But introducing the ideal numbers $\sqrt{2}$, $(1+\sqrt{5i})/\sqrt{2}$, $(1-\sqrt{5i})/\sqrt{2}$ rescues the unique factorisation:

$$6=2\cdot 3=\sqrt{2}\cdot\sqrt{2}\cdot\left[(1+\sqrt{5i})/\sqrt{2}\right]\cdot\left[(1-\sqrt{5i})/\sqrt{2}\right].$$

$$6=(1+\sqrt{5i})(1-\sqrt{5i})=\sqrt{2}\cdot\left[(1+\sqrt{5i})/\sqrt{2}\right]\cdot\sqrt{2}\cdot\left[(1-\sqrt{5i})/\sqrt{2}\right].$$

Again, both factorisations coincide.

Kummer dedicated himself to the intense study of this new group of cyclotomic integers, conveniently extending these new ideal numbers. He managed to demonstrate that for a special type of prime number, the so-called regular prime numbers, all of the demonstration's reasoning is recuperated and the last theory is demonstrated. Then, he studied regular prime numbers and demonstrated that there are only three non-regular prime numbers smaller than 100: 37, 59 and 67. He dealt with these cases last, until he had demonstrated the theorem for all exponents lower than 100.

The Académies de Sciences, encouraged by these new advances, decided to give the subject a decisive boost, and in 1850 it again offered a prize to anyone who

could resolve the last theorem definitively, once and for all. The board of judges was formed by Augustin-Louis Cauchy, Joseph Liouville, Gabriel Lamé, Joseph-Louis François Bertrand and Michel Chasles. Once all the time periods and conceded extensions had passed, Cauchy wrote: "Eleven documents have been submitted to the Board. But none of them has resolved the proposed problem. The Board, however, has noted that the work registered under number 2 contained a new solution to the problem in the special case developed by Fermat himself, in other words, when exponent equals 4.

"Therefore, having increased the prize on many occasions, the understanding of the subject was stalled where Kummer had left off. However, the mathematical sciences should congratulate themselves for the work carried out by geometricians, with their desire to resolve the problem, especially by Kummer; and the Board thinks that the Académie should take the honourable and useful decision of, leaving the matter of the competition aside, awarding the medal to Mr. Kummer, for his marvellous research on complex numbers formed by roots of the unit and integers."

Thus, the prize was awarded to Kummer in 1857, when he had not even entered the competition! In this gesture, deep recognition of his work was implicit. His contributions were of such magnitude that they opened the doors to new and vast fields of research: regular prime numbers, the theory of ideals, cyclotomic numbers, numbers of class, theory of bodies of class, and many more new ideas and concepts began to be researched.

The last theorem had given rise to a spectacular advance in mathematics, but continued to be impenetrable. After more than two hundred years, progress was as follows. The first case had been demonstrated for a large number of exponents, all of which complied with the conditions established by Germain and Legendre. The general case for the four exponents $n = 3, 4, 5, 7$ had also been demonstrated. But even so, there were still many remaining. The last theorem, with all its fascination, was starting to be a thorn in the side for many mathematicians.

A question of genus

In 1908, German businessman and mathematician Paul Wolfskehl established a prize of 100,000 German marks (approximately one million pounds in today's money) for anyone who was capable of proving Fermat's theorem. The maximum, non-extendable deadline for resolving the problem was 13 September 2007. Probably, Wolfskehl thought that 99 years was a reasonable period of time to resolve the problem.

SOLUTIONS FOR THE FIRST 100 PRIME EXPONENTS P

p	First case	General case
3	Fermat (c. 1638)? Euler (c. 1760)	
4	Fermat (c. 1638)	
5	Germain (1823)	Dirichlet & Legendre (c. 1825)
7	Legendre (1823)	Lamé (1839)
11	Germain (1823)	Kummer (c.1850), *p* regular
13	Legendre (1823)	Kummer (c.1850) , *p* regular
17	Legendre (1823)	Kummer (c.1850) , *p* regular
19	Legendre (1823)	Kummer (c.1850) , *p* regular
23	Germain (1823)	Kummer (c.1850) , *p* regular
29	Germain (1823)	Kummer (c.1850) , *p* regular
31	Legendre (1823)	Kummer (c.1850) , *p* regular
37	Legendre (1823)	Kummer (c.1857) , *p* irregular
41	Germain (1823)	Kummer (c.1850) , *p* regular
43	Legendre (1823)	Kummer (c.1850) , *p* regular
47	Legendre (1823)	Kummer (c.1850) , *p* regular
53	Germain (1823)	Kummer (c.1850) , *p* regular
59	Legendre (1823)	Kummer (c.1857) , *p* irregular
61	Legendre (1823)	Kummer (c.1850) , *p* regular
67	Legendre (1823)	Kummer (c.1857) , *p* irregular
71	Legendre (1823)	Kummer (c.1850) , *p* regular
73	Legendre (1823)	Kummer (c.1850) , *p* regular
79	Legendre (1823)	Kummer (c.1850) , *p* regular
83	Germain (1823)	Kummer (c.1850) , *p* regular
89	Germain (1823)	Kummer (c.1850) , *p* regular
97	Legendre (1823)	Kummer (c.1850) , *p* regular

A considerable number of mathematicians made great efforts to extend the list of exponents for which the last theorem was verified, both for the first and general cases. Occasionally these advances were achieved by simply polishing criteria obtained previously or perfecting calculation methods, or sometimes they led to the discovery of new lines of research. In 1909, Wieferich demonstrated that if there was a solution for the first case of Fermat's theorem, then $2^{p-1} - 1$ had to be a multiple of p^2. This condition turned out to be very powerful. In fact, at the time no prime number had been found that verifies it. Until in 1913 Meissner found the case $p =$ 1,903, and in 1922 Beeger found the case $p =$ 3,511. In 1910 Mirianoff extended Wieferich's results and demonstrated that if there was a solution for the first case of

Fermat's theory then $3^{p-1} - 1$ would also have to be a multiple of p^2. This meant that the cases of $p = 1{,}903$ and $p = 3{,}511$ were covered. In 1971, thanks to the help of computers, Brillhart, Tonascia and Weinberger investigated all prime numbers up to $3 \cdot 10^9$, without finding any more that verified Wieferich's condition and, therefore, demonstrating the first case of Fermat's theorem up to that exponent. In the following years the quantity of prime numbers under study increased, and by 1990 all exponents of Fermat's theorem up to $2{,}327 \cdot 10^{19}$ had been demonstrated.

Regarding the general case, Vandiver studied Kummer's work in depth, and in 1929 provided a whole series of criteria that had to by met by irregular prime numbers in order to verify Fermat's last theorem. In 1954 the same Vandiver, now with the help of computers, verified the exponents where $p < 2{,}521$. Twenty years later the list had increased to $p < 4{,}000{,}000$. Amid this frantic progress to refine the criteria and perfect the calculations, the mathematical community received a pleasant surprise.

In 1922, Briton Louis Mordell (1888-1972) had conjectured that in all algebraic curves of genus greater than 1 there was only a finite number of rational points. The genus of an algebraic curve would become a measure of its complexity. Curves of

TO FERMAT, FOR SAVING MY LIFE

There are several theories on why Wolfskehl decided to set up his Fermat prize. When he was young, he was stricken by multiple sclerosis and he had to abandon the practice of medicine for a more relaxed activity: mathematics. Some sources say that, at the point of committing suicide due to a love affair, his study of the theory convinced him that the beauty of mathematics was greater than that of any woman. Therefore, in a very literal sense, Fermat had saved his life. Others cite a more prosaic reason: Wolfskehl had used the legacy to reduce the sum that would be inherited by his unfaithful wife.

German mathematician Paul Wolfskehl.

genus zero would be the simplest, and as the genus increases, so does the complexity of its individual points. In 1983, German Gerd Faltings (1954–) received a Fields Medal for the demonstration of this conjecture, which brought with it a completely new tool for solving Fermat's equation. For exponent $n = 2$, the curve $x^2 + y^2 = z^2$ turned out to be of genus 0, and has infinite Pythagorean triples as solutions. But for $n > 2$ the curve $x^n + y^n = z^n$ is of genus greater than 1! From this result if followed that if there were solutions to the last theorem, their number would be finite. The mathematical community was convinced that the final proof was close, and that it would arrive through the door opened by Mordell and Faltings. But they were wrong.

A bridge between two worlds

At the end of the 1980s experts knew of a wide range of conjectures which, if they were proven, would also prove the last theorem by implication, at least for the case of large exponents: the *abc* conjecture, the Szpiro, Vojta and Bogomolov-Miyaoka-Yau conjectures, etc. To everyone's surprise, this exclusive club was to include a new member: the Taniyama-Shimura conjecture.

Proposed in the 50s and refined in the 70s, the Taniyama-Shimura conjecture postulates an extraordinary and unexpected link between two families of mathematical objects which at a glance appear to be completely unrelated: elliptic curves, intimately related to cubic equations of the type Diophantus studied in his day, and modular forms, developed at the end of the 19th century by Frenchman Jules Henri Poincaré. Its announcement is owed to the contributions of Japanese mathematicians Goro Shimura (1930–) and Yutaka Taniyama (1927-1958). Both young men were well-known and worked together in Tokyo, in the desolate, post-war Japan, and theirs is a beautiful story of intellectual complicity, over whose scientific success a shadow of eventual tragedy is cast.

THE *ABC* CONJECTURE

This conjecture was presented in 1985 Joseph Oesterlé and David Masser. In simplified form it states the following: given integer co-primes, a, b, c, such that a + b = c, and designating d as the product of the different prime factors of a, b and c, it holds that d will rarely be less than c.

The first world: elliptic curves

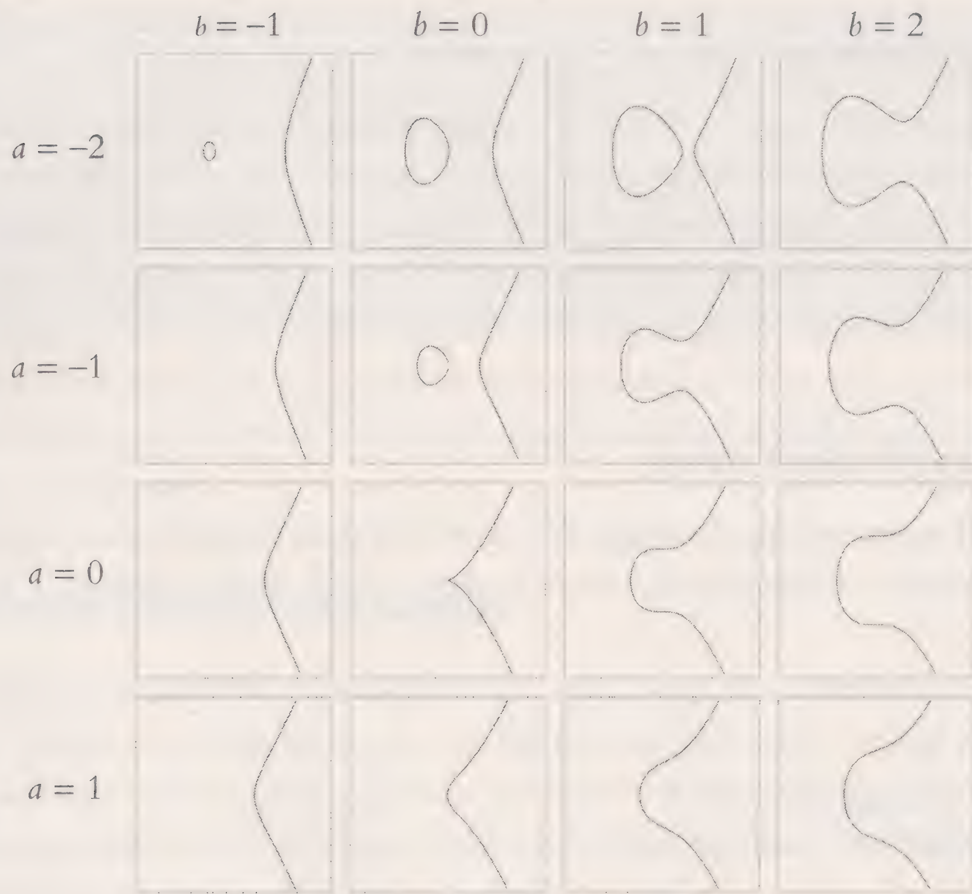
One way of approximating the length of a curve is to draw straight lines which join the finite number of points on the curve, as shown in the diagram below:

As the size of the segments is reduced, the sum of their lengths moves closer and closer to that of the curve, a process that is known as polygonal approximation. In the case of some curves, there is a value L which is the maximum limit of all possible polygonal approximation; it is then said that the curve has an arc length L. From the study of arc lengths of ellipses came what are called elliptic functions, and from them, elliptic curves.

German Karl Theodor Wilhelm Weierstrass (1815–1897) demonstrated that all elliptic curves were defined by a cubic curve of the type:

$$y^2 = x^3 + ax^2 + bx + c,$$

where a, b and c are real numbers. For $c = 0$ and different values of a and b, the elliptic curves offer the suggestive appearance shown on the next page:



Elliptic curves for c = 0 *and different values of* a *and* b.

An important challenge in number theory, already faced by our old friend Diophantus, is to find integer solutions to an equation of this type. For example, the following cubic equation:

$$y^2 = x^3 - 2$$

can also be written as:

$$x^3 - y^2 = 2.$$

A positive integer solution to this equation is equal to finding that natural number or numbers are exactly 'in the middle' of the cube and the square of any two other natural numbers. Well, the first mathematician in history to provide an answer to this was... Pierre de Fermat, who demonstrated that 26 is the only number that meets that condition. That is to say $x^3 = 27$ and $y^2 = 25$ and therefore the only positive integer solutions to the above cubic equation are $x = 3$ and $y = 5$. To summarise this wonderful network of connections between the main players of our story,

German mathematician Karl Theodor Wilhelm Weierstrass, who made important contributions in the field of elliptic curves, in a painting by Conrad Fehr.

we should remember here that one of the modern mathematical tools employed for the study of elliptic curves is the Iwasawa theory, the subject of Andrew Wiles' doctoral thesis. Wiles would rightly say that "in a certain sense, all the mathematics that I have produced follow a trail left by Fermat."

Finding the solutions to an elliptic equation is, in most cases, virtually impossible, which is why mathematicians tend to study them in the framework of 'limited' spaces of numbers called 'modulos'. To understand this concept we can imagine the way in which the hours of the day are conceived. If, for example, we talk about an event that happened 30 hours after midnight on any given day, it is understood that the event took place at 6 in the morning (of the second day, in this case). Mentally, we have counted 24 whole hours (one day), we have moved from midnight on one day to the next and then added the difference $30 - 24 = 6$, to determine the specific time of the event. In mathematical terms we say that the time of day obeys an arithmetic modulo 24 (due to the number of hours in a day) which verifies, as we have seen, that $30 \equiv 6$. If, instead of 30 hours, we had been talking about 38, the event would have taken place at 14:00 and, therefore, in the arithmetic modulo 24 it is said that $38 \equiv 14$ (and, similarly,

ELLIPTIC CURVES AND CRYPTOGRAPHY

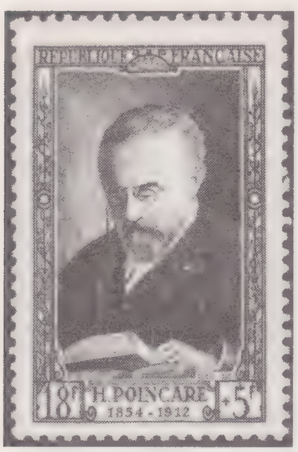
There are certain mathematical operations which, once carried out, are difficult to reverse, for example, finding the prime factors of a large integer number. The RSA algorithm is one of the pillars of modern cryptography and is based on the previous operation to construct codes that are theoretically impossible to break. Another operation considered to be 'irreversible' is finding the discrete logarithm of an elliptic curve. Therefore, in 2009, the US government started using certain encryption algorithms based on this property as standard for the transmission of ultra-secret information.

that $24 \equiv 0$). Regardless of the number of hours that have passed since any event, be it 36 or 36,000, the calculation of the time of day at which it took place at can only be taken as a value between 0 and 23. This type of arithmetic allows common adding, subtracting, multiplying and dividing operations to be carried out, and the result of any operation of this kind in an arithmetic modulo 24 will be, again, just one of the 24 numbers from 0 to 23. Now let's go back to the elliptic equations. What solutions could these equations have in, for example, modulo 2? A maximum of 4, the following:

$$x = 0, y = 0$$
$$x = 0, y = 1$$
$$x = 1, y = 0$$
$$x = 1, y = 1$$

By making use of the powerful tool that is modular arithmetic, mathematicians can characterise an elliptic equation, not by means of its 'absolute' solutions, which are difficult to find, but using the number of its solutions for each module. So, any elliptic equation would be defined by a series E of infinite elements E_1, E_2, E_3... where the value of each element is the number of solutions of the equation in modulo 1, 2, 3... An equation with two solutions in modulo 2, for example (0,0) and (1,0), would have a value of $E_2 = 2$ in its series.

The second world: modular forms



French stamp dedicated to mathematician Jules Henri Poincaré.

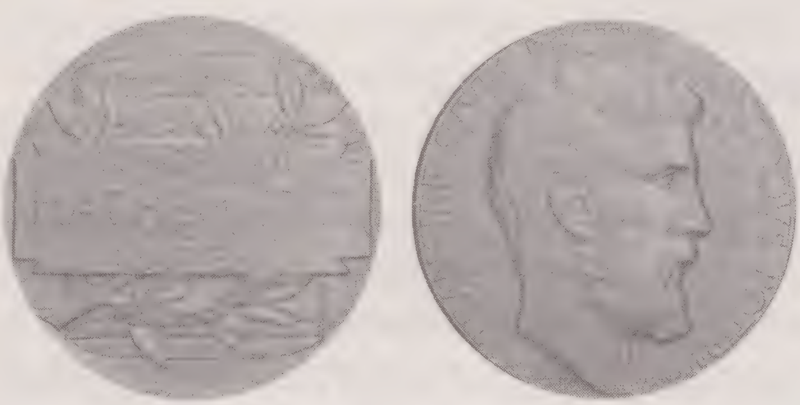
Modular forms are in large part a creature of Jules Henri Poincaré, one of the most notable scientific minds in history. Some of his work in mathematical physics can be considered direct antecedents of Einstein's theory of relativity. He was also an important publicist and philosopher of science. Poincaré was the last mathematician to have in-depth knowledge of all the disciplines of his time, an achievement which, today, is impossible due to the enormous variety of sectors covered by mathematics. Poincaré, who was already a recognised mathematician in his adolescence, had, like Euler and Gauss, a photographic memory, especially in the field of spatial memory. That perhaps explains his position as the founder

of topology, the branch of mathematics that studies the spatial properties of objects that remain unchanged during certain deformations. Topology is a kingdom where symmetry reigns, and few mathematical objects have such profound and extensive symmetry as that of modular forms.

To imagine a modular form, in the sense of creating a visual idea of it, is completely impossible. Suffice it to say that they inhabit a four-dimensional space, a space governed by a geometry that is very different to that which we are accustomed. In our daily

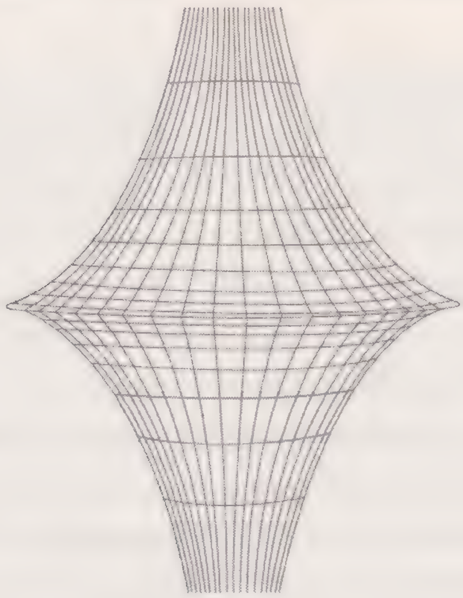
POINCARÉ'S 'LAST THEOREM'

Although none of them have achieved the notoriety or legend of Fermat's last theorem, there are a handful of conjectures in mathematics the resolution of which would be considered a genuinely historic event. Nobody would argue that they include the Goldbach conjecture and, *primus inter pares*, Riemann's hypothesis, both included in the field of number theory. Others are the P versus NP problem, an key question for computing, and in topology, the so-called 'Poincaré hypothesis'. To everyone's surprise, between 2002 and 2003 Russian Grigori Perelman made public a diagramatical proof of this conjecture which in 2006, thanks to the complementary work of other specialists, was unofficially declared valid. Perelman, a mathematician whose brilliance was matched by his eccentricity, rejected a Fields Medal that was awarded to him that year for his astonishing achievement and, claiming that the academic world is littered with dishonesty, completely abandoned mathematics shortly afterwards. As in the case of the other problems included by the Clay Institute in its 1999 list of the seven 'Millennium Prize Problems', proof of the Poincaré hypothesis is rewarded with one million dollars. At the time of the publication of this challenge, Perelman's proof had still not been published in any authorised magazine and its validity had not been certified by a committee of experts established for that purpose. Therefore, according to the framework of the prize, Perelman had no right to it. It should be pointed out that Perelman has never claimed it either.



The Fields Medal, which Perelman rejected for his work on the 'Poincaré Hypothesis'.

life we can get by quite well assuming, as did Euclid, that only one parallel line passes through a point outside another straight line. Since the 19th century, however, it has been known that this is not necessarily true, but a simple convention. Other alternative geometries can be defined, all of them coherent, in which there are no parallel lines or, on the other hand, an infinity of them. The latter case is known as 'hyperbolic geometry', the plane of which when represented in two dimensions adopts suggestive forms such as the following:



Parallel lines from the point of view of hyperbolic geometry.

It is in this peculiar hyperbolic world that modular forms are housed, and where the most extraordinary symmetries are on show, as if we were dealing with extremely rare orchids. To define any of these 'mathematical orchids', mathematicians use what is called their M series, which consists of an infinite series of elements M_1, M_2, M_3..., each of which corresponds to a numerical value that indicates the amount of the 'ingredient' 1, 2, 3... that it contains.

The bridge: the Taniyama-Shimura conjecture

In the mid-1950s, Japan was still struggling to recover from the traumas of the Second World War. Although the economy was growing little by little, living conditions continued to be hard. Academia did not escape from the dearth of facilities and the scarcity of paid research positions forced students to compete fiercely for them. If the area of interest of the researcher was a field that was little associated with productive activity, the situation became even more

difficult, and even the most enthusiastic of candidates was turned away from studying pure mathematics.

This was not the case with young Yutaka Taniyama, the eighth son of a provincial doctor, who despite the desperate conditions and countless health problems suffered during his infancy, had the strength to move to the capital city, with barely any resources, to sign up to the university and concentrate on his mathematical studies. In 1954 he made friends with a brilliant colleague, Goro Shimura, who was one year older than him, and with whom he often met up in cheap cafés or restaurants to discuss advanced concepts of number theory, the speciality that most attracted both of them. There could not have been a greater contrast between their personalities: Taniyama was absent-minded to the point of being chaotic; he worked remarkably hard, often at night, and showed so little interest in things unmathematical that he was bordering on eccentric. Shimura, on the other hand, got up to work at dawn, was organised and meticulous. While his colleague wore the same suit made of a strange metallic material day, in day out and as a rule never tied his shoe laces, Shimura maintained a healthy interest in his appearance and related more naturally to the rest of his colleagues.

What they did have in common was a desire to bury themselves in the most recent developments from the international mathematical scene, the reason why they decided, in 1955, to organise a symposium on number theory to which they invited reputed mathematicians from all ove the world. Of the 36 problems presented to the symposium's attendees, four of them had been proposed by Taniyama, and they put forward, in a notoriously vague manner, certain types of relationships between modular forms, a subject that at the time had been somewhat forgotten by specialists, and Diphantine equations. Taniyama had noticed that some of the elements of the *E* series of some elliptic equations corresponded exactly with the elements of the *M* series of certain modular forms, although at this early stage of development he was unable to present the grounds of this interesting coincidence.

This and other issues were discussed at the symposium, and according to some sources, one of the attendees was illustrious French mathematician André Weil, who, in an informal conversation with Taniyama suggested to him the idea that his intuitions pointed towards a profound and general relationship between modular forms and elliptic equations. This version of events, which was later demonstrated to be, to put it lightly, imprecise, survived the passing of time to the point where the Taniyama-Shimura conjecture was known for a while as the Shimura-Weil conjecture, or the Taniyama-Shimura-Weil conjecture, a serious case of erroneous attribution which

required intervention from the American Serge Lang, quite a few years later, to put it right.

However Taniyama's initial intuition, expressed as it was in such non-specific terms, provoked little interest. The only one who believed in that intuition right from the beginning was his faithful friend Shimura, and for the next two years both of them worked side-by-side to develop it more solidly. In 1957 Shimura was invited to join Princeton, where he thought he would be able to exchange ideas with reputed specialists and thus continue to advance the matter, but a tragic event was to jeopardise the ambitious project. On 17 November of the same year Taniyama decided to end his life. The suicide note read: "Until yesterday I had no definite intention of killing myself. [...] As to the cause of my suicide, I don't quite understand it myself, but it is not the result of a particular incident, nor of a specific matter. Merely may I say, I am in the frame of mind that I lost confidence in my future. [...] At any rate, I cannot deny that this is a kind of betrayal, but please excuse it as my last act in my own way, as I have been going my own way all my life." Taniyama was 35.

His friend's tragic end only reinforced Shimura's will to finish their joint work and thus, in a sense, honour the memory of a great mathematician. After several years of work, Shimura was shaping a conjecture that stated, in very simple terms, that all elliptic curves were modular, a conjecture which with time would become known as the Taniyama-Shimura conjecture. In the words of American mathematician Barry Mazur (who we will talk about later) it is a "wonderful conjecture... but to begin with it was ignored because it was so ahead of its time. When it was first proposed it was not taken up because it was so astounding. On the one hand you have the elliptic world, and on the other you have the modular world. Both these branches of mathematics have been studied intensively but separately. Then along comes the Taniyama-Shimura conjecture which is the grand surmise that there's a bridge between these two completely different worlds. Mathematicians love to build bridges..."

Taniyama did not live to see his great intuition become one of the most beautiful results in modern mathematics, but his name and that of his colleague and friend Shimura now have a place of honour in the history of the discipline, and what would undoubtedly have been even more surprising and satisfactory is that their work would become a fundamental chapter in the demonstration of the most celebrated theorem in number theory and in mathematics in general.

The epsilon conjecture

In the eyes of the mathematical community, Taniyama-Shimura and Fermat's last theorem had nothing more in common that the simple property of being conjectures. But finding unexpected relationships between seemingly unrelated concepts, as we have seen, is one of the essential goals of mathematics. In this specific case, the unexpected connection was to be made by a German mathematician, Gerhard Frey. Frey, a specialist in number theory, was fascinated by the relationships between this field and algebraic geometry, of which the Taniyama-Shimura conjecture is a magnificent example. In 1978 he read and was very impressed by the work of American Barry Mazur in which concepts such as modularity and elliptic curves were interrelated, and the connection of which he took pains to make more explicit. (Mazur's seminal article in this field was called *Modular Curves and the Eisenstein Ideal*, and among its most enthusiastic readers were Ken Ribet and Andrew Wiles). Frey started to develop a surprising idea that he tried to define during a stay of several weeks at Harvard, the university where Mazur was a professor. Finally, in 1984, at a mathematics conferences held in the German town of Oberwolfach, Frey postulated a hypothesis that would open up a new and revolutionary path for approaching the last theorem.

The German mathematician proceeded as follows: Given any solution to the theorem, for example: $a^p + b^p = c^p$, there is an associated elliptic curve of the form $y^2 = x(x - a^p)(x + b^p)$, where a, b and c are co-prime integers greater than zero, and p is a prime number greater than 2. This curve belongs to a special group of elliptic curves (subsequently called 'Frey curves'), which have a very interesting property – they are not modular. Now, the Taniyama-Shimura conjecture stated that all elliptic curves are modular, from which it follows that if Taniyama-Shimura is true, an 'aberration' such as Frey's curve, or, a curve that is both elliptic and not modular, is impossible. Consequently, if Frey's conjecture was correct, given that all possible solutions to the last theorem would in turn be a Frey curve, the Taniyama-Shimura affirmation that such curves do not exist is the same as saying that the last theorem has no solutions and, therefore... that it is true! This unsuspected link between one conjecture and another, as we will see, was to be the helping hand that would allow Wiles to reach the summit of his quest.

Although Frey's ideas were enormously interesting, the truth is that his conjecture was not sufficiently concrete for other mathematicians to be able to verify it. A little 'mathematical muscle' was needed to give the German's intuitions the definitive form of a conjecture. And in the context of mathematical muscle in the last 75 years

French mathematician Jean-Pierre Serre at the Abel Prize ceremony, which took place on 3 June 2003. (Photograph courtesy of the Abel Institute.)

it is difficult to overlook Frenchman Jean Pierre Serre (born in 1926), one of the only mathematicians (the other is John Griggs Thompson) to receive the two most prestigious honours in the discipline: the Fields Medal (1954), which he received at a record young age of 27 years, and the Abel prize (2003); this is an achievement of a similar magnitude to that of winning two Nobel prizes. Serre, who incidentally in his youth attended the symposium organised in 1955 by Taniyama and Shimura, was interested in Frey's formula and wrote a personal note in this regard to his colleague and compatriot Jean François Mestre, a note that he later converted into an article. In this document the French mathematician formalised Frey's hypothesis in somewhat different terms (filling in its gaps by means the introduction of the so-called 'modular Galois representations'), and he officially elevated it to the category of a conjecture. The epsilon conjecture, as it was informally named, if it was correct, implied a necessary relationship between Taniyama-Shimura and the last theorem: if the first is true, the other cannot be, and vice versa.

THE HAND THAT ROCKS THE CRADLE

American Barry Mazur, born in 1937, is one of the outstanding figures in the field of number theory in recent decades. His article *Modular Curves and the Eisenstein Ideal* is one of the principal reasons for the renewed interest in modularity as seen in the work of young mathematicians such as Frey, Ribet and Wiles and, therefore, of the actual solution to the Fermat's problem. Mazur had characterised number theory as a field of mathematics "which produces, without effort, innumerable problems that have a sweet, innocent air about them, tempting flowers; and yet these flowers swarm with bugs, waiting to bite the tempted flower-lovers who, once bitten, are inspired to excesses of effort."

GERHARD FREY, MATHEMATICIAN AND CRYPTANALYST

Frey was born in 1944, in the German area of Tübingen. He studied physics and mathematics at the city's university. His speciality is number theory, and among his most relevant contributions, apart from the epsilon conjecture, is the method known as 'Weil descent' for the resolution of elliptic curves on finite bodies. This proof put an end to a promising field of cryptography.

From conjecture to theorem

The implications of the epsilon conjecture were such that all specialists in number theory tried their luck with it. Among them was a brilliant young American mathematician, Kenneth Ribet, professor at Berkeley in the mid 1980s. Ribet was a student of Mazur's at Harvard, where he took his doctorate, and he inherited from his teacher a fascination for links between number theory and geometric algebra which Kummer illuminated in his day and which would offer the context in which all those involved in the solution to the last theorem operated. Ribet applied himself to the task of demonstrating the epsilon conjecture, and one day he saw the light. In his own words:

> "I was completely enthralled. I just sort of wandered back to my apartment in a cloud, and I sat down and I ran through my argument, and it worked. It really worked. And at the conference [the International Mathematics Conference held at the University of Berkeley, San Francisco, in 1986], I started telling a few people that I'd done this, and soon, large groups of people knew, and they were running up to me, and they said, 'Is it true that you've proved the epsilon conjecture?' And I had to think for a minute, and all of a sudden, I said, 'Yes. I have.'"

This simple, sincere confession allows us to get an idea of what might go through the head of a mathematician when he realises that he has rescued an iota of truth from an ocean of ignorance, truth with a capital T, because if anything characterises mathematics, it is the illumination of truths in the deepest and most absolute sense

of the term. Ribet himself later commented that in his doctoral years, paraphrasing Gauss, he had said of the last theorem: "It is one of those problems on which we can no longer say anything useful." At that time Ribet had no idea that his work would be fundamental when it came to making those words true a few years later. *The epsilon conjecture* passed into the pages of the history books and was replaced by Ribet's theorem. The last theorem could now be attacked with 'state-of-the-art' mathematical weapons.

So, now what?

The cards were on the table. Whoever proved the Taniyama-Shimura conjecture would prove Fermat's last theorem. Easy to say, but difficult, extraordinarily difficult, to do. In the end, nearly forty years had passed since the symposium in which Taniyama had put forward his original idea, and still no one had made any significant steps towards proving the conjecture. In fact, the overwhelming majority of number theory specialists thought it highly improbable that the conjecture would be proven within a few decades (remember Mazur's characterisation in this regard: "A wonderful conjecture... but to begin with it was ignored because was so ahead of its time.") A large part of the problem lay in the fact that both modular forms and elliptic curves, the two types of mathematical objects which the conjecture stated were interrelated, were infinite. Whoever attempted to take on the enormous task of proving the Taniyama-Shimura conjecture was going to have to resolve not only this great problem, but a multitude of other smaller, but equally difficult ones, a process during which the slightest error at the time of calibrating the promise of one particular line of attack could lead to years and years of useless work. If the theorem is considered the Everest of mathematics it could be said that Taniyama, Shimura, Mazur, Frey, Serre and Ribet had found a new route to the peak, which had been previously hidden, but this route was still highly treacherous.

Chapter 6

The Proof

$x^n + y^n = z^n$: no solutions.
I have discovered a truly remarkable proof of this,
but I can't write it now because my train is coming.
Graffiti found in a New York subway station, 1988

It was an afternoon in the summer of 1986 and Andrew Wiles was at the house of a friend drinking iced tea. As they talked, his friend let slip that Ribet had proved the epsilon conjecture. This statement triggered a deluge of emotions in Wiles. In his own words, "I knew that moment the course of my life was changing because this meant that to prove Fermat's last theorem, I just had to prove the Taniyama-Shimura conjecture. From that moment, that was what I was working on."

Wiles abandoned the other projects in which he was currently involved and dedicated his heart and soul to solving the problem. It was a decision that would leave him for all intents and purposes cut off from the world for the next seven years, although as he himself would explain years later, he knew that he enjoyed a significant advantage: nobody else had any idea how to tackle the problem. However this relative lack of competition also had its drawbacks: "I realised after a while that talking to people casually about Fermat was impossible, because it just generates too much interest, and you can't really focus yourself for years unless you have the kind of undivided concentration that too many spectators would have destroyed." Wiles hardly suspected that later along the path to glory, he would have to work not just against the clock, but also under the scrutiny of the entire world.

The boy who dreamed of proving Fermat's last theorem

One of the most common anecdotes about Wiles involves his early fascination with the last theorem after having heard the story at the age of twelve when he stumbled across it in a book on popular mathematics. To the young Wiles, the idea of the lost proof was akin to stories of buried treasure and adventure in far off lands and

THE DANGERS OF EXTRAPOLATION

As mentioned in the previous chapter, in the 1990s the Last Theorem had been verified for exponents of up to four million. If it was correct for exponents of this size, why were mathematicians still so determined to prove the result for all possible exponents? Was it not highly unlikely, indeed almost impossible, that an inconceivably high exponent would appear to contradict the formula, as if suddenly falling from the sky with no warning at all? Was the mathematical community not perhaps being overly stringent? Psychological matters aside, for a conjecture that affirms a property for an infinite number of cases, a collection of 'experimental' results, however large it may be, can be considered as nothing more than an indication, never a proof. And mathematics is built on proofs, or rather on indisputable truths, and herein lies not only its magic, but its power as a scientific tool. Furthermore, the history of number theory is plagued with counterexamples to hypotheses, or should we say, 'empirical indications'. As an example, in his time Euler conjectured that the following equation had no solution:

$$x^4 + y^4 + z^4 = w^4.$$

The circuits of many a computer have been worn out over the course of decades trying to provide a counterexample to this hypothesis but to no avail. The temptation to affirm that Euler's conjecture was correct for all cases was strong, until in 1988 Noam Elkies astonished the mathematical community:

$$2{,}682{,}440^4 + 15{,}365{,}639^4 + 187{,}960^4 = 20{,}615{,}673^4.$$

And this was not all: not only did Elkies find a solution, but he also showed there to be an infinite number of solutions. Computers, it would seem, have their limitations.

treacherous caverns. He was so fascinated that he began to tackle the problem with no greater tools than school-level arithmetic. This anecdote shows the extent to which Fermat's seemingly simple statement is even within the reach of a child, yet those who tackle it are soon drawn deeper and deeper into the unutterably complex depths of the last theorem. Inevitably the young Wiles had to give up his attempts but Fermat's spell would remain with him for life.

Although he was born in Cambridge in 1953, Andrew John Wiles graduated in mathematics from Oxford, where his father, Maurice Frank Wiles, had been a professor of theology. However he studied for his PhD at Cambridge under the supervision of the Australian John Coates. His PhD thesis focused on the arithmetic of elliptic curves based on methods from what is referred to as 'Iwasawa theory'. At the start

of the 1980s, Wiles accepted a teaching post at Princeton University and became an editor of the *Annals of Mathematics,* a prestigious journal. It seemed that Wiles had forgotten all about his early fascination with Fermat. As he himself would confess later on, "It's not that I forgot about it; it was always there. I always remembered it, but I realised the only techniques we had to tackle it had been around for 130 years, and it didn't seem they were really getting to the root of the problem. So, when I went to Cambridge, my advisor, John Coates, was working on Iwasawa theory and elliptic curves, and I started working with him." One of the beautiful symmetries in our story is that this theory would turn out to become one of the keys to his proof of the last theorem. Then in 1986 Ribet proved the epsilon conjecture, and Wiles was once again overwhelmed by his old fascination.

Counting infinities

During the following seven years, Wiles worked obsessively on his proof. For the first two, he concentrated exclusively on immersing himself in the problem and exploring it from all possible points of view in the search for a plan of attack that would allow the proof to work. (In this respect, the Briton John Edensor Littlewood remarked on one occasion that a mathematician should know a problem "as one knows one's own tongue rolling it round the inside of the mouth".) The attic on the third floor of the Wiles's family home on the outskirts of Princeton became the most active room in the house, although this was perhaps difficult to believe given how little noise emerged from it. Wiles unplugged the telephone and, never having liked computers much, worked exclusively on paper, filling thousands and thousands of sheets with all kinds of formulae alongside occasional drawings, diagrams and graphs. Progress was extremely slow. On occasion, he would attempt to make use of tools that were already known to allow him to go from one step of the proof to another, to gently massage it to make it fit his purposes. There were also times when he simply tried to invent new mathematics. Throughout the initial phase of the process, Wiles maintained a strict silence about what he was working on.

To begin with, he considered the possibility of 'counting' each of the elliptic functions (the number of which, let's remember, is infinite) on the one hand, and modular elliptic functions (also infinite) on the other, and proving that each count was the same in both cases. This strategy proved to be ineffective, although during the course of his study Wiles stumbled on an important result that would simplify

Andrew Wiles in 2000. (Photograph by C.J. Mozzochi, Princeton New Jersey.)

his task. Instead of having to prove the Taniyama-Shimura conjecture for all elliptic curves, he realised that he only needed to do so for a subgroup of these curves, the semi-stable ones.

In this aspect of his work, Wiles sought inspiration from Galois theory, named in honour of its inventor, the ill-fated French mathematician Évariste Galois (1811–1832). Galois, a genuinely tragic figure in the history of mathematics, had the brilliant idea of studying the possible solutions (or roots) of a polynomial in terms of the permutations that can be based on these roots. This idea was subsequently developed by Augustin Louis Cauchy and Arthur Cayley with extraordinary success. For example, the second-degree polynomial

$$x^2 - 4x + 1 = 0$$

has roots $x_1 = 2 + \sqrt{3}$ and $x_2 = 2 - \sqrt{3}$.

Both roots satisfy the following algebraic equations:

$$x_1 + x_2 = 4,$$
$$x_1 \cdot x_2 = 1.$$

Both equations continue to hold if we swap x_1 and x_2:

$$x_2 + x_1 = 4,$$
$$x_2 \cdot x_1 = 1.$$

Galois conducted in-depth research into functions that were invariant with respect to the permutations of their roots, and based on this work, defined what is known as the Galois group of an equation. Thus, as an example, the Galois group of the polynomial $x^2 - 4x + 1 = 0$ consists of two permutations – the identity permutation (which leaves the roots 'as they are') and the transposition permutation (as shown in the example above).

The properties of Galois groups are an extremely powerful tool that makes it possible to characterise highly complex structures, and Wiles made good use of them when it came to overcoming the first great obstacle in his proof. Specifically, he started to define elliptic equations in terms of their Galois representations and proved that these could be associated with certain characteristic elements of modular shapes. This allowed Wiles to reformulate the 'counting' problem in manageable terms. This first and fundamental advance would, on its own, have been worthy of recognition by the mathematical community. However it was a means to an end and after having completed this first step, Wiles had already been working intensely for two years.

Working in the most complete isolation, where did Wiles find the strength to continue? In his own words, "I was concentrating very hard, and I found that being with young children is the best possible way to relax. When you're talking to young children, they simply aren't interested in Fermat, at least at this age. They want to hear a children's story, and they're not going to let you do anything else." Perhaps

THE ANGRY GENIUS

A portrait of Évariste Galois aged 15 drawn by one of his classmates.

Évariste Galois was a young and passionate Frenchman who had not hesitated to embrace the republican cause during the turbulent years of Louis Philippe I, the last king of France. He is also one of the most remarkable geniuses from the history of mathematics. His hot temper was poorly suited to the rigours of academic research, and his inability to pass the entrance examinations for college meant that his mathematical writings went almost unnoticed among his contemporaries. Galois overcame his academic frustration with increasing political radicalism which led to him being challenged to a duel by an artillery officer who sympathised with the monarchy. Aware of his disadvantage when it came to the handling of arms, the night before the duel he wrote a feverish letter to his acquaintance Auguste Chevalier, a distinguished mathematician, summarising his scientific ideas. The following morning he received a fatal wound to the abdomen and died a few hours later. The letter he had written to Chevalier contained the seed of the theory that would later be named after him, one of the fundamental pillars of modern algebra. Galois was 21 years old.

Wiles was lucky his children did not show the precocious interest in Fermat that he himself had developed as a child.

The new counting method devised by Wiles also showed interesting parallels with the material he had studied for his PhD under the supervision of Coates – Iwasawa theory. It was 1988, and he could already feel the breath of other mathematicians on his neck. Imagine then, the look on his face when on 8 March he read in the headlines of the major newspapers that Fermat's last theorem had been proved by someone from Japan, Yoichi Miyaoka. Although details of the proof had not yet been made public, certain experts in the field had publicly expressed their confidence in the general strategy of the argument presented by the young Japanese mathematician. However, a few months later, it became clear that the supposed proof contained a fatal flaw. The mansion remained in darkness; Wiles still had work to do.

Flach, Katz and flickers of light

However, to the great disappointment of Wiles, Iwasawa theory did not prove as useful as he had hoped. His frustration was obvious when he said:

MIYAOKA'S PROOF OF FERMAT'S LAST THEOREM

Robert Langlands at the 61st birthday of the mathematician Pierre Deligne in Princeton. The ceremony took place in 2006 (fotografía: C.J. Mozzochi, Princeton N.J.).

The failed proof of Fermat's last theorem suggested by Miyaoka was based on an approximation of the problem derived from the philosophy of parallelism. Inspired by certain general principles exposed at the end of the 1960s by the Canadian mathematician Robert Langlands in the so-called 'Langlands programme', this philosophy sought to tackle a number of problems from number theory by making use of tools from algebraic geometry. One example of the successful use of the theory is the proof of the Mordell conjecture by the skilful German mathematician Gerd Faltings. It was also Faltings who, as one of the experts responsible for confirming the proof was valid, revealed the specific error committed by the Japanese mathematician. However Miyaoka's tireless efforts to fix the problem never led to its solution.

> "I really believed that I was on the right track, but that did not mean that I would necessarily reach my goal. It could be that the methods needed ... would not be invented for a hundred years. So even if I was on the right track, I could be living in the wrong century."

After five years of isolation, Wiles decided to come out and make contact with certain colleagues, including his old mentor John Coates. Coates spoke highly of the work of one of his students, Matthias Flach, who, based on the work of the Russian Victor Kolyvagin, appeared to have developed a powerful tool that would make it possible to simplify the elliptic equations that continued to resist Wiles' efforts. In the words of Wiles, it seemed to be "purpose built". All that remained was to extend the partial results of Flach-Kolyvagin to cover the scope of his particular problem, and Wiles immediately set to the task with renewed vigour. After a number of months of intense work, the new technique appeared to be delivering the expected results, yet Wiles could not help but feel a certain insecurity for having based his complex proof on a technique that had been discovered so recently, and about which he could not help but harbour certain doubts. The time had come for him to abandon his secrecy and convert his individual struggle into a modest conspiracy.

The figure selected by Wiles as both confidante and aid was a colleague from the university, an expert in the type of algebra used in the Flach-Kolyvagin work. Nick Katz tells of the moment when Wiles confided in him about the nature of the project he had been working on for the last six years: "January of 1993. Andrew came up to me one day at tea, asked me if I could come up to his office; there was something he wanted to talk to me about. I had no idea what this could be. I went up to his office. He closed the door. He said he thought he would be able to prove Taniyama-Shimura. I was just amazed. This was fantastic."

Aside from his mathematical knowledge, Wiles turned to Katz because he was sure that he would keep his secret. And he was right. Nonetheless, there remained the challenge of designing a working method that, regardless of the hours they would need to spend discussing and reviewing equations, would not arouse the suspicions of other colleagues at the university. Wiles and Katz had a most sibylline idea. The former announced a new class for PhD students entitled "Calculations on Elliptic Curves", open to both undergraduates and faculty members. Its content would be based on nothing else but a step-by-step explanation of Wiles' proof, so that Katz, in his capacity as assistant, would be able to scrutinise it without arousing suspicion. The few PhD students who registered were not long in dropping out because the

ENLIGHTENMENT

In the midst of his struggle to apply Iwasawa theory to his proof, Wiles decided to take a trip to a lake near the campus where he hoped to be able to relax and to allow, in his own words, his "sub-conscious to work on him". The belief that the unconscious will continue to work on solving a problem is a common aspect of the creative experience, especially among mathematicians. The Frenchman, Henri Poincaré, left us with a vivid portrait of a moment of enlightenment that resulted from a similar mental process in his discovery that Fuchs functions (later renamed as automorphic functions) were identical to those from non-Euclidean geometry: "Just at this time I left Caen, to go on a geologic excursion... The changes of travel made me forget my mathematical work. We entered an omnibus to go some place or other. At the moment when I put my foot on the step the idea came to me, without anything in my former thoughts seeming to have paved the way for it. On my return to Caen, for conscience's sake, I verified the result at my leisure."

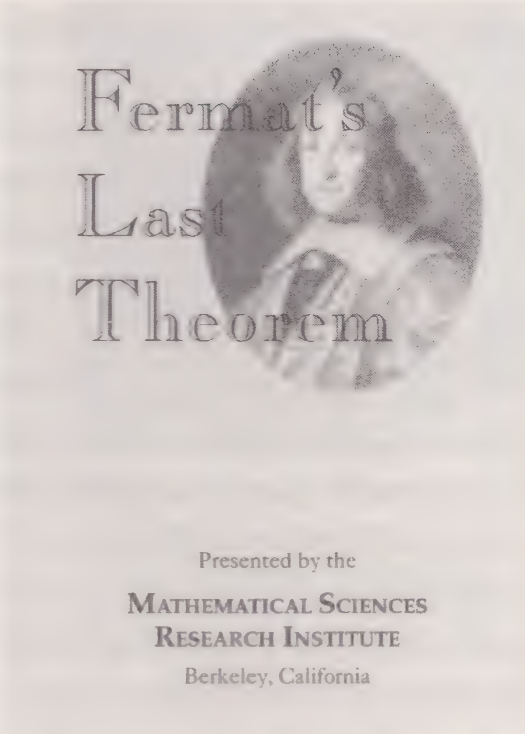
classes discussed such esoteric material. "It's impossible to follow stuff if you don't know what it's for... It's pretty hard even if you do know what it's for... But after a few weeks, I was the only guy in the audience," Katz recalls.

This arrangement bore fruit, and Katz was unable to find any problems with the steps of Wiles' proof. To be even more certain, Wiles invited a third member to join the conspiracy, Peter Sarnak, another of his colleagues from Princeton. "I think I'm about to prove Fermat's last theorem," Wiles confided in an astonished Sarnak. "I remember that night finding it quite difficult to sleep," the latter reported.

However one obstacle still remained. Certain elliptic curves continued to resist his technique. It was then that the figure of Mazur emerged, when one of his articles gave Wiles the idea of changing a parameter in his work. Wiles remembers:

> "And, I just kept working out the details, and time went by, and I forgot to go down to lunch, and it got to about tea-time, and I went down and Nada [Wiles' wife] was very surprised that I'd arrived so late, and then I told her that I believed I'd solved Fermat's last theorem. I was convinced that I had Fermat in my hands, and there was a conference in Cambridge organised by my advisor, John Coates. I thought that would be a wonderful place. It's my old home town, and I'd been a graduate student there."

The Cambridge conference was scheduled to take place between 21 and 23 June, and Wiles worked at full pace to organise that last seven years' of labour. The manuscript for the final version of the proof ran into 200 pages and was ready just in time for him to catch the plane to the United Kingdom.



Sleeve from the video of Fermat's Last Theorem. The video was recorded in 1993 and numerous figures from the world of mathematics, such as Andrew Wiles and Ken Ribet took part.

An early morning email

In 1993, the British mathematician John Horton Conway was the brightest talent in the mathematics department at Princeton University. A renowned expert in set theory, game theory and geometry, he had also invented one of the first cellular automata, a self-modifying computer program known as 'the game of life'. On 23 June, Conway, always an early riser, had been first to enter the mathematics department. A number of weeks had now passed since one of his colleagues, Andrew Wiles, had left for a conference in Cambridge, and Conway, an active figure in the international mathematical community, had been hearing rumours associated with a significant, impending result, despite knowing nothing of its nature. With the first light of dawn illuminating the mountains of papers and books crammed into every last corner of his office, he switched on his computer and read the emails he had received during the preceding night. One of the last to download, written at 5:53 had the following concise subject line: "Wiles proves F.L.T."

"I'm still not satisfied, Andrew"

Wiles returned to Princeton that Friday, emotionally exhausted. "I haven't let go of this problem for nearly seven years," Wiles confessed, "It's gone on and on day to day. I've almost forgotten the experience of getting up and thinking about something else." Wiles dealt with the deluge of messages wishing him congratulations, some of which were from people thanking him for allowing them to have seen a

solution to the problem in their lifetime. The repercussions of the event were so large that, in an almost unprecedented step, the US magazine *People* listed Wiles as one of the 25 most fascinating figures of the year.

Although Wiles's proof was still the object of documentaries and television programmes, the academic world began a much less rewarding, but equally necessary, process. The proof was peer reviewed by a committee of experts, an essential step for it to be accepted as correct. For a document as complex as Wiles's proof, the final version of which ran to around 200 pages, this process would go on for a number of months. Although on more than one occasion, the same process had revealed fatal flaws in proofs that were at first believed to be correct (one example being Miyaoka's proof five years earlier), there were few people who believed that the revision would be anything more than a formality, given the previous scrupulous revision to which Wiles had subjected the paper. At the same time, neither did anyone expect the document to be faultless. It was normal for small errors to be detected during the review, but in the majority of cases, these did not affect the main argument and were easily rectified.

Wiles had decided to publish his proof in the academic journal *Inventiones Mathematicae*, whose editor was none other then Barry Mazur. Mazur appointed a team of experts for the review, which included names such as Gerd Faltings and Nick Katz. The latter dedicated almost every day of July and August to combing through Wiles's work, line by line, specifically chapter 3 which was 70 pages long. The revision continued the same routine, day in day out. If Katz had a doubt about any part of the steps of the proof, he would email Wiles who would reply as soon as possible, always managing to clear up the doubt to his full satisfaction. Except in one case.

Katz had checked around two thirds of the chapter he had been allocated and encountered difficulties in understanding a specific argument, whose validity required the establishment and application of an abstruse mathematical tool known as the 'Euler system', which was precisely one of the elements imported from the Flach-Kolyvagin methods that both Katz and Wiles had checked together during the sessions of their phoney course. It was a particularly complicated part of the proof and, instead of sending an email, Katz faxed his concerns to Wiles. Wiles replied with his usual speed, however Katz, unlike on previous occasions, was not satisfied and re-sent the question with the innocent remark "I'm still not satisfied, Andrew." Both men exchanged faxes again, to no success. In September, Wiles had no alternative but to accept that something was not right.

The mathematician Nick Katz, who was first to work with Wiles on his research and was later one of the reviewers of the proof. (Photograph: C.J. Mozzochi, Princeton New Jersey.)

At first, he insisted on trying to solve the problem by making various changes to the Flach-Kolyvagin system. He went back to his custom of shutting himself away in his attic and working in complete isolation. Yet try as he might, he was unable to solve the logical fissure that prevented him from obtaining the Euler system he so badly needed. The pressure was starting to get to him. Katz recalls: "In October, the only people who were aware of the existence of the error were myself, Illusie [Luc Illusie, the French mathematician whom Katz contacted in July to help with the revision], the reviewers of the other chapters and Andrew. My attitude was as to be expected from any reviewer: strict secrecy." Although the whole world expected that the assessment of such a manuscript would take a number of months, half way through the autumn of 1993, the mathematical community began to suspect that there was a serious problem and began to show its impatience. The inboxes of departmental email accounts were inundated with all sorts of speculation. (Simon Singh cites one such email, dated 18 November by Joseph Lipman, from the University of Purdue, which is especially edifying about the state of mind among specialists. "There are many rumours buzzing around about one or more gaps in Wiles's proof. Does gap mean crack, fissure, crevasse, chasm or abyss?") In the end, after the phenomenal announcement made at the conference in June, no one except from the reviewers had been able to cast eyes on the official version of the proof. The pressure on Wiles was starting to become considerable, and some journalists began to approach experts about the state of British mathematician and of the proof itself.

Towards the end of November, the problem had still not been solved. This led Wiles to post the following message, dated 4 December, on the sci.math newsgroup:

> "In view of the speculation on the status of my work on the Taniyama-Shimura conjecture and Fermat's Last Theorem, I will give a brief account of the situation. During the review process a number of problems emerged, most of which have been resolved, but one in particular I have not yet

AN ENORMOUS CARPET

In a television programme broadcast in 2000 on the story of Wiles and his proof of the last theorem, Peter Sarnak made the following comment about Wiles's efforts to correct the error in his original proof: "Every time he would try and fix it in one corner, it would sort of – some other difficulty would add up in another corner. It was like he was trying to put a carpet in a room where the carpet was bigger than the floor, but he could put it in in any corner, and then when he ran to the other corners, it would pop out there. And whether you could not fit the carpet in the room was not something that he was able to decide."

> settled... I believe that I will be able to finish this in the near future using the ideas explained in my Cambridge lectures. The fact that a lot of work remains to be done on the manuscript makes it still unsuitable for release as a preprint. In my course in Princeton beginning in February I will give a full account of this work.
>
> Andrew Wiles."

The situation was extremely uncomfortable for all those involved, especially for Wiles, who would later claim:

> "The first seven years I'd worked on this problem, I loved every minute of it. However hard it had been, there'd been setbacks often, there'd been things that had seemed insurmountable, but it was a kind of private and very personal battle I was engaged in. And then, once there was a problem with it, doing mathematics in that kind of rather over-exposed way is certainly not my style, and I have no wish to repeat it."

On Sarnak's advice, Wiles obtained the help of a former student, Richard Taylor, a bright young mathematician. Both set to work, however notwithstanding their considerable efforts, most of 1994 passed without them finding a way to adjust the Flach-Kolyvagin method to the requirements of the original proof..

Revelation

That same autumn, Wiles, disheartened and dejected, having reached the limits of his strength, decided to throw in the towel once and for all. There was no human way to reconstruct the proof. However, a simple professional hunch caused him to go back three years and return to studying the Flach-Kolyvagin technique from the start, in order to at least be able to determine why a line of attack that had seemed so promising had ended in failure. Once again, he sat at the desk in his office, the same one that had seen both his rise to glory and his subsequent fall from grace. It was a peaceful Monday morning, 19 September, a day whose most intricate details Wiles will remember forever:

> "And I was sitting here at this desk just seeing exactly what the problem was, when suddenly, totally unexpectedly, I had this incredible revelation. I realised what was holding me up was exactly what would resolve the problem I had had in my Iwasawa theory attempt three years earlier. It was the most – the most important moment of my working life. It was so indescribably beautiful; it was so simple and so elegant."

Wiles sat for twenty minutes looking at the documents, astonished and in disbelief, tears falling from his eyes.

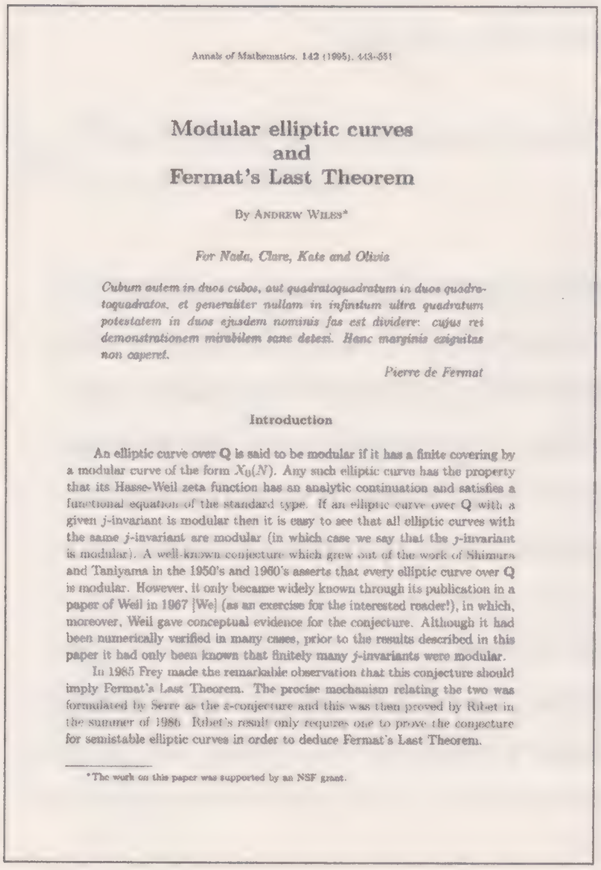
> "Then, during the day, I walked around the department. I'd keep coming back to my desk and looking to see if it was still there. It was still there. Almost what seemed to be stopping the method of Flach and Kolyvagin was exactly what would flatten the Iwasawa theory. My original approach to the problem from three years before would make it work. So, out of the ashes seemed to rise the true answer to the problem. So, the first night, I went back and slept on it. I checked through it again the next morning, and by eleven o'clock, I was satisfied and I went down and told my wife, 'I've got it. I think I've got it. I've found it.' And it was so unexpected, I think she thought I was talking about a child's toy or something and said, 'Got what?' And I said, 'I've fixed my proof. I've got it.'"

Nada's birthday was on 3 October and, although it was a few days early, her husband had given her what was certainly an extraordinary gift. During the following days, Taylor and Wiles reviewed the details of the new corrected proof, without

finding any errors. A few days later, two manuscripts were published. The first, extremely long and signed by Andrew Wiles, was entitled "Modular Elliptic Curves and Fermat's Last Theroem"; the other, shorter version was entitled "Ring-Theoretic Properties of Certain Hecke Algebras", and was signed by Wiles and Richard Taylor. The first contained the proof of the Taniyama-Shimura conjecture for semi-stable elliptic curves, with one of their crucial steps being supported by the details of the second. Both received considerable commentary and were submitted for publication in the *Annals of Mathematics*, which, after the appropriate peer review, approved the findings and published them in the May 1995 issue.

The medal he never received

At last the whole of the mathematical community could breathe a sigh of relief. Wiles's extraordinary achievement, when almost nobody expected that the drama would have a happy ending, was felt throughout the breadth and depth of the scientific world. Academia did not hesitate to recognise his work with all types of awards, such as the 1995 Wolf prize, one of the most prestigious and well-endowed in the discipline; the Schock prize for the same year; the Royal Society Medal, and the Ostrowski prize in 1996; the Cole prize for number theory in 1997, which had previously been won by Goro Shimura, Barry Mazur and Karl Rubin; and, naturally, the Fermat prize, established in 1989 to award contributions in the fields in which the French mathematician had distinguished himself.



Annals of Mathematics, 142 (1995), 443-551

Modular elliptic curves
and
Fermat's Last Theorem

By ANDREW WILES*

For Nada, Clare, Kate and Olivia

Cubum autem in duos cubos, aut quadratoquadratum in duos quadratoquadratos, et generaliter nullam in infinitum ultra quadratum potestatem in duos ejusdem nominis fas est dividere: cujus rei demonstrationem mirabilem sane detexi. Hanc marginis exiguitas non caperet.

Pierre de Fermat

Introduction

An elliptic curve over **Q** is said to be modular if it has a finite covering by a modular curve of the form $X_0(N)$. Any such elliptic curve has the property that its Hasse-Weil zeta function has an analytic continuation and satisfies a functional equation of the standard type. If an elliptic curve over **Q** with a given j-invariant is modular then it is easy to see that all elliptic curves with the same j-invariant are modular (in which case we say that the j-invariant is modular). A well-known conjecture which grew out of the work of Shimura and Taniyama in the 1950's and 1960's asserts that every elliptic curve over **Q** is modular. However, it only became widely known through its publication in a paper of Weil in 1967 [We] (as an exercise for the interested reader!), in which, moreover, Weil gave conceptual evidence for the conjecture. Although it had been numerically verified in many cases, prior to the results described in this paper it had only been known that finitely many j-invariants were modular.

In 1985 Frey made the remarkable observation that this conjecture should imply Fermat's Last Theorem. The precise mechanism relating the two was formulated by Serre as the ε-conjecture and this was then proved by Ribet in the summer of 1986. Ribet's result only requires one to prove the conjecture for semistable elliptic curves in order to deduce Fermat's Last Theorem.

*The work on this paper was supported by an NSF grant.

The first page of the paper Modular Elliptic Curves and Fermat's Last Theorem, *published by Wiles in the journal* Annals of Mathematics.

In 1998 he also received the Faisal prize, and in 1999, the Clay Institute award. And we shouldn't forget to mention the Wolfskehl prize, the value of which had reduced dramatically, largely as a result of the hyperinflation

that struck Germany in the 1930s, but which still had a very reasonable value of £30,000.

However, the string of honours received by Wiles is above all notable for one glaring omission, the Fields Medal. This medal, which is awarded once every four years, was established in 1936 by the Canadian mathematician John Charles Fields. It is awarded at the International Mathematical Union (IMU) Conference and, despite having a decidedly modest prize of only $10,000, is doubtless the most prestigious award in the discipline. Fields wanted to offer support to young mathematicians, meaning that the minimum age for the award had been fixed at 40. However, and especially since many claim that creative activity in this particular discipline reaches its peak in the third decade of life, the Fields medal is considered a fair prize and an accurate rating of the most distinguished contributions to mathematics. Unfortunately for both the British mathematician and the medal itself, Wiles was already 41 by the time of the 1994 ceremony. However the IMU could not go without recognising his contribution and for the first time in its history, it made an extraordinary award of a silver plate in acknowledgement of Wiles's exceptional discovery.

Epilogue. Life after Fermat?

In the introduction to his celebrated conference held in 1900 discussing the state of mathematics at the start of the new century, the German mathematician David Hilbert wrote: "Who of us would not be glad to lift the veil behind which the future lies hidden; to cast a glance at the next advances of our science and at the secrets of its development during future centuries?" And with respect to the last theorem, he added: "The attempt to prove this impossibility offers a striking example of the inspiring effect that such a very special and apparently unimportant problem may have upon science. For Kummer, incited by Fermat's problem, was led to the introduction of ideal numbers and to the discovery of the law of the unique decomposition of the numbers of a circular field into ideal prime factors – a law that today stands at the centre of the modern theory of numbers and the significance of which extends far beyond the boundaries of number theory into the realm of algebra and the theory of functions."

Hilbert penned his reflections decades before the contributions made by Mordell, Taniyama-Shimura and Frey, and most certainly without any knowledge at all of the form the proof would take for Wiles. Who can guess what extraordinary

advances will be derived from the work of these figures? With the magical bridge between elliptic curves and modular forms suggested by Taniyama-Shimura, what other unexpected links could arise in the future between other regions on the great mathematical map, now considered to be far apart?

However the importance of the last theorem does not lie in its scientific relevance, great though this may be, but in the catalytic effect that a discovery of this scale can have for generations of future researchers. For centuries, the challenge of Fermat stood like an unconquerable citadel, with the spears of many a mathematician being deflected by its walls. The faith and conviction shown by Wiles in tackling the solution to such a unique problem on his own has doubtless inspired others to dedicate their talents to other open problems that might at first glance seem impossible.

What does Wiles himself have to say in this respect? The natural reserve of the British mathematician does not leave room for any grand pronunciation regarding these matters. However, it would be impossible not to conclude this book in his own words, this time spoken after the definitive acceptance of his second proof, the crowning jewel of a lifetime's dream:

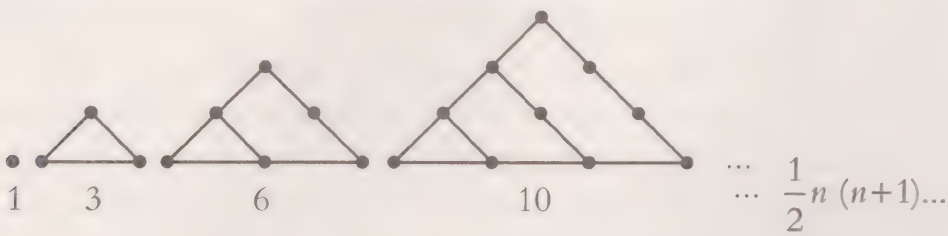
> "I had this rare privilege of being able to pursue in my adult life, what had been my childhood dream. I know it's a rare privilege, but if one can really tackle something in adult life that means that much to you, then it's more rewarding than anything I can imagine. There is a certain sense of sadness, but at the same time there is this tremendous sense of achievement. There's also a sense of freedom. I was so obsessed by this problem that I was thinking about it all the time – when I woke up in the morning, when I went to sleep at night – and that went on for eight years. That's a long time to think about one thing. That particular odyssey is now over. My mind is now at rest."

Appendix

Polygonal Numbers

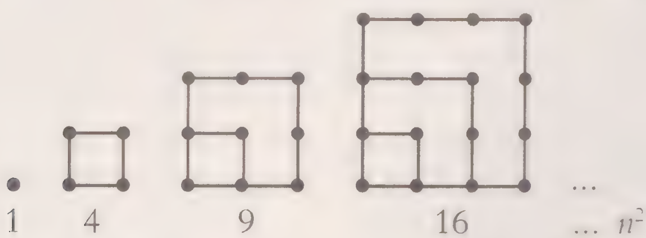
A polygonal number is one that can be represented as points arranged in a regular polygon, starting at 1. These numbers have attracted the attention of mathematicians from time immemorial. The Greeks attributed magical properties to them, related to the shapes with which they were associated, and Diophantus dedicated a treatise to their study.

The triangular numbers are based on equilateral triangles:



The series obtained is 1, 3, 6, 10, 15, 21, 28, 36, 45, 55, 66, 78, 91, 105… and is given by the formula beside the illustration.

The square numbers are based on regular quadrilaterals – squares



The series obtained is 1, 4, 9, 16, 25, 36, 49, 64, 81, 100, 121, 144, 169, 196, 225… and is given by the formula n^2.

The pentagonal numbers are based on pentagons:

The series obtained is 1, 5, 12, 22, 35, 51, 70, 92, 117, 145, 176, 210, 247, 287, 330… and is given by the formula beside the illustration.

In the same way, it is possible to obtain the hexagonal, numbers, heptagonal ones, and so on…

Fermat was the first to realise that all the natural numbers can be expressed as a sum of at most three triangular numbers, four square numbers, five pentagonal numbers, etc. The following tables verifies this for triangular and pentagonal numbers.

n	Sum of triangular numbers	Sum of square numbers	Sum of pentagonal numbers
1	= 1	$= 1^2$	= 1
2	= 1 + 1	$= 1^2 + 1^2$	= 1 + 1
3	= 3	$= 1^2 + 1^2 + 1^2$	= 1 + 1 + 1
4	= 3 + 1	$= 2^2$	= 1 + 1 + 1 + 1
5	= 3 + 1 + 1	$= 2^2 + 1^2$	= 5
6	= 6	$= 2^2 + 1^2 + 1^2$	= 5 + 1
7	= 6 + 1	$= 2^2 + 1^2 + 1^2 + 1^2$	= 5 + 1 + 1
8	= 6 + 1 + 1	$= 2^2 + 2^2$	= 5 + 1 + 1 + 1
9	= 6 + 3	$= 3^2$	= 5 + 1 + 1 + 1 + 1
10	= 10	$= 3^2 + 1^2$	= 5 + 5
11	= 10 + 1	$= 3^2 + 1^2 + 1^2$	= 5 + 5 + 1
12	= 6 + 6	$= 2^2 + 2^2 + 2^2$	= 12
13	= 10 + 3	$= 3^2 + 2^2$	= 12 + 1
14	= 10 + 3 + 1	$= 3^2 + 2^2 + 1^2$	= 12 + 1 + 1
15	= 15	$= 3^2 + 2^2 + 1^2 + 1^2$	= 5 + 5 + 5
16	= 15 + 1	$= 4^2$	= 5 + 5 + 5 + 1
17	= 10 + 6 + 1	$= 4^2 + 1^2$	= 12 + 5
18	= 15 + 3	$= 3^2 + 3^2$	= 12 + 5 + 1
19	= 10 + 6 + 3	$= 3^2 + 3^2 + 1^2$	= 12 + 5 + 1 + 1
20	= 10 + 10	$= 4^2 + 2^2$	= 5 + 5 + 5 + 5
21	= 21	$= 4^2 + 2^2 + 1^2$	= 5 + 5 + 5 + 5 + 1
22	= 21 + 1	$= 3^2 + 3^2 + 2^2$	= 22
23	= 10 + 10 + 3	$= 3^2 + 3^2 + 2^2 + 1^2$	= 22 + 1
24	= 21 + 3	$= 4^2 + 2^2 + 2^2$	= 12 + 12
25	= 15 + 10	$= 5^2$	= 12 + 12 + 1
26	= 15 + 10 + 1	$= 5^2 + 1^2$	= 12 + 12 + 1 + 1
27	= 21 + 6	$= 5^2 + 1^2 + 1^2$	= 22 + 5
28	= 28	$= 5^2 + 1^2 + 1^2 + 1^2$	= 22 + 5 + 1
29	= 28 + 1	$= 5^2 + 2^2$	= 12 + 12 + 5
30	= 15 + 15	$= 5^2 + 2^2 + 1^2$	= 12 + 12 + 5 + 1

Bibliography.

ACZEL, A.D., *El último teorema de Fermat*, Mexico, Fondo de Cultura Económica, 2004. [Spanish popular science book on Fermat's Last Theorem.]

GHEVERGHESE, G.J., *The Crest of the Peacock: Non-European Roots of Mathematics*, London & New York : I.B. Tauris, 1991.

MAHONEY, M.S., *The mathematical Career of Pierre de Fermat, 1601-1665* , Princeton University Press, 1994.

RIBENBOIM, P., *Fermat's Last Theorem for Amateurs*, New York, Springer Verlag, 1999.

SINGH, S., *Fermat's Last Theorem: the Story of a Riddle that Confounded the World's Greatest Minds for 358 Years*, London: Fourth Estate, 1997.

Index

Fermat's Enigma

The quest to prove Fermat's last theorem

No other conjecture in the history of mathematics has caused such widespread debate as that stated by the brilliant French mathematician Pierre de Fermat in 1637. The simplicity of its formulation contrasts with the great mathematical depths to which its study leads, and the quest to prove it introduces us to some extraordinary mathematical minds.